

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

**CHARACTERIZATION OF TWO, DESICCATION LINKED,
GROUP 1 LEA PROTEINS FROM THE RESURRECTION
PLANT *XEROPHYTA HUMILIS***

ZEKARIAS GEBREMEDHIN GINBOT

Thesis Presented for the Degree of

DOCTOR OF PHILOSOPHY

in the

Department of Molecular and Cell Biology,

Faculty of Science

University of Cape Town

May 2011

Declaration

I declare that this submission is my own work and that, to the best of my knowledge, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgment has been made in the text.

Signed _____ at the University of Cape Town on the _____ of May, 2011.

Acknowledgements

I thank my supervisors, Professors Jill M. Farrant and Nicola Illing for their professional guidance throughout this work. I also would like to acknowledge Professor Farrant's financial support at times when I did not have other source of funding. Thanks to Keren Kooper and Johana Nell, colleagues in the lab, for the friendship I enjoyed during this PhD.

I thank my parents Gebremedhin and Jewet and aunts Letebrehan and Futru for all the sacrifices they made to send me to school despite they themselves have never been to school. My special thanks go to the Parish of St. Michael's Catholic Church, in Cape Town, especially Fr Harrie Hovers, the parish priest, for the assistance provided to me and my family. I also thank the Sigrid Rausing Trust for providing funding during the last two years of my PhD. Being a refugee, a student and parent at the same time, the last four years have been very difficult for me and my family. So, I thank my own family; my wife Lete and my children Shenhet, Lurdi and Gedna for enduring all the trouble as a result of the choices I made in pursuing my studies.

Abbreviations

ABA	Absciscic acid
AWC	Absolute water content
CD	Circular Dichroism
cDNA	Complementary DNA
CDNN:	CD Spectra Deconvolution software
Cfu:	Colony forming units
CS	Citrate Synthase
ELISA	Enzyme-Linked Immunosorbent Assay
ERD	Early Responsive to Dehydration
FTIR	Fourier Transfer Infrared
GST	Glutathione S-transferase
IPTG	Isopropyl β -D-thiogalactoside
<i>kDa</i>	Kilo Daltons
LB	Luria-Bertani
LDH	Lactate dehydrogenase
LEA	Late Embryogenesis Abundant
MALDI	Matrix-assisted laser desorption/ionization
mRNA	Messenger RNA
MS	Mass spectrometry
MWCO	Molecular weight cut off
NMR	Nuclear Magnetic Resonance or
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline

PCF	Pair correlation function
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PMSF	Phenylmethanesulfonyl fluoride
pNPP	Para-nitrophenyl phosphate
PONDR	Prediction of Naturally Disordered Regions
POPP	Protein or Oligonucleotide Probability Profile
PSV	protein storage vacuole
PVDF	Polyvinylidene fluoride
RT-qPCR	Reverse Transcription quantitative Polymerase Chain Reaction
RWC	Relative water content
SD	Standard deviation
TCA	Trichloroacetic Acid
TEM	Transmission electron microscope
TFE	Trifluoroethylene
UTR	Untranslated region
UV	Ultra violet
YT	Yeast extract and Tryptone

Tables and Figures

A. Figures

Figure 2.1 Predicted amino acid sequences of <i>XhLEA1-4S1</i> (A) and <i>XhLEA1-1S2</i> (B).	33
Figure 2.2 Sequence alignments of <i>XhLEA1-4S1</i> and <i>XhLEA1-1S2</i> against	34
Figure 2.3 Insertion of <i>Bam</i> HI and <i>Eco</i> RIenzyme restriction sites into LEA ..	35
Figure 2.4 Colony screening by PCR for <i>pGEX-XhLEA1-4S1</i> (lanes 3-6) and <i>pGEX</i> -.....	36
Figure 2.5 Colony screening for <i>pGEX-XhLEA1-4S1</i> (A) and <i>pGEX-XhLEA1-1S2</i> (B).	36
Figure 2.6 Optimization of <i>XhLEA1-4S1</i> (A) and <i>XhLEA1-1S2</i> (B) fusion protein	37
Figure 2.7 Purification of GST- <i>XhLEA1-4S1</i> and GST- <i>XhLEA1-1S2</i> fusion proteins.	38
Figure 2.8 MS spectra of GST+XhLE1-4S1 (A and B) fusion protein, m/z= 800 - 1600. ...	40
Figure 2.9 MS spectra of GST+XhLE1-1S2 (A and B) fusion protein, m/z= 800 - 1600....	41
Figure 2.10 Checking the specificity and cross reactivity of anti <i>XhLEA1-4S1</i> (A) and	43
Figure 2.11 Checking the specificity of <i>XhLEA1-4S1</i> (A) and <i>XhLEA1-1S2</i> (B)	44
Figure 2.12 Western blot expression studies of <i>XhLEA1-4S1</i> (A) and <i>XhLEA1-1S2</i> (B)	45
Figure 2.13 Western blot expression study of <i>XhLEA1-4S1</i> in dry roots. In 'A' (Ponceau...	46
Figure 3.1 Immunolocalization of XhLEA1-4S1 in hydrated leaf.....	60
Figure 3.2 Immunolocalization of XhLEA1-4S1 in dry leaf.....	61
Figure 3.3 Immunolocalization of XhLEA1-1S2 in hydrated leaf.....	62
Figure 3.4 Immunolocalization of XhLEA1-1S2 in dry leaf.....	63
Figure 3.5 Immunolocalization of XhLEA1-14S in hydrated root	64
Figure 3.6 Immunolocalization of XhLEA1-14S in dry root	65
Figure 3.7 Immunolocalization of XhLEA1-1S2 in hydrated root	66
Figure 3.8 Immunolocalization of XhLEA1-1S2 in dry root	67
Figure 3.9 Immunolocalisation of XhLEA1-4S2 protein in seeds of <i>X. humilis</i>	73
Figure 3.10 Immunolocalisation of XhLEA1-1S2 protein in seeds of <i>X. humilis</i>	74
Figure 3.11 Test for clustering of gold labels for <i>XhLEA1-4S1</i> (top left) and	77

Figure 4.1 Insertion of <i>Bam</i> HI and <i>Eco</i> RI restriction sites into <i>XhLEA1-4S1</i> and	89
Figure 4.2 Colony-PCR screening of clones grown after transformation with reco	89
Figure 4.3 Expression of <i>XhLEA1-4S1</i> and <i>XhLEA1-1S2</i> recombinant proteins in.....	90
Figure 4.4 Confirmation of identities of <i>XhLEA1-4S1</i> and <i>XhLEA1-1S2</i>	91
Figure 4.5 Thermal stability test of lysates containing <i>XhLEA1-4S1</i> and.....	92
Figure 4.6 Purification of <i>XhLEA1-4S1</i> and <i>XhLEA1-1S2</i> recombinant proteins	92
Figure 4.7 Prediction of structural disorder of <i>XhLEA1-4S1</i> (A) and <i>XhLEA1-1S2</i>	93
Figure 4.8 A Uversky plots of <i>XhLEA1-4S1</i> (A) and <i>XhLEA1-1S2</i> (B) amino acid.....	94
Figure 4.9 CD spectrophotometer readings of <i>XhLEA1-4S1</i> (A), <i>XhLEA1-1S2</i> (B)	95
Figure 4.10. CS aggregation in the presence or absence of <i>XhLEA1-4S1</i> (A),	99
Figure 4.11 Effect of XhLEA proteins on the activity of 0.12mg desiccated CS enzyme...100	

B. Tables

Table 1.1 Nomenclature of LEA groups (Harada <i>et al.</i> , 1989;	7
Table 1.2 Non-plant LEA proteins with expression evidence	12
Table 2.1 Primers used for the amplification of the full-length <i>XhLEA1-4S1</i> and.....	26
Table 2.2 Amino acid composition of <i>XhLEA1-4S1</i> and <i>XhLEA1-1S2</i> proteins.....	34
Table 2.3 Collection of IgG antibody serum fractions from DE52 column	42.
Table 3.1 LEA proteins with known subcellular localisation	54
Table 3.2 Analysis of gold label distribution in hydrated and dry <i>X. humilis</i> leaf cells using anti- <i>XhLEA1-4S1</i> primary antibody.....	68
Table 3.3 Analysis of gold label distribution in hydrated and dry <i>X. humilis</i> leaf cells using anti- <i>XhLEA1-1S2</i> primary antibody.....	69
Table 3.4 Analysis of gold label distribution in hydrated and dry <i>X. humilis</i> root cells using anti- <i>XhLEA1-4S1</i> primary antibody.....	71

Table 3.5 Analysis of gold label distribution in hydrated and dry <i>X. humilis</i> root cells using anti- <i>XhLEA1-1S2</i> primary antibodies.....	71
Table 3.6 Analysis of gold label distribution patterns for <i>XhLEA1-4S1</i> and <i>XhLEA1-1S2</i> proteins in the cell wall-plasma membrane compartment of fully hydrated and dry <i>X. humilis</i> leaf and root tissues by Pair correlation function, $g(r)$	75
Table 3.7 Analysis of gold label distribution patterns for <i>XhLEA1-4S1</i> and <i>XhLEA1-1S2</i> proteins in protein storage vacuoles (PSVs) of <i>X. humilis</i> by Pair correlation function, $g(r)$	76
Table 4.1 Primers used for cloning and sequencing of <i>XhLEA1-4S1</i> and <i>XhLEA1-1S2</i>	84
Table 4.2 <i>XhLEA1-4S1</i> and <i>XhLEA1-1S2</i> aggregation test. A two-cycle vacuum	97

Abstract

Studies on resurrection plants and other anhydrobiotic organisms show expression of Late Embryogenesis Abundant (LEA) proteins associated with desiccation tolerance. However, the precise role of these proteins has not been described. This study was undertaken to investigate expression, structure and function of *XhLEA1-4S1* and *XhLEA1-1S2*, Group 1 LEA proteins from *Xerophyta humilis*, in order to shed light on their role in desiccation tolerance. Complementary DNA (cDNA) of these XhLEAs were cloned into bacterial expression vectors and the recombinant proteins expressed in *E. coli*. Antibodies were generated and used in determination of expression conditions and immunolocalization studies.

Western blot analysis showed expression of both XhLEA proteins in dry leaves, roots and seeds of *X. humilis*. No expression was detected in the hydrated/re-hydrated tissues. Expression in leaves was induced during drying at relative water contents (RWCs) of $\leq 55\%$. Subcellular localization studies showed the association of these proteins with the plasma membrane associated with cell wall in root and leaf tissue and in seeds, with protein storage vacuoles (PSVs). *XhLEA1-4S1* and *XhLEA1-1S2* are shown to be unstructured in solution, highly hydrophilic with PONDR disorder values of 98% and 88% and corresponding hydropathy scores of 0.3398 and 0.3983 respectively. Circular Dichroism (CD) studies on these XhLEAs showed increased α -helical content on addition of 50% trifluoroethylene (TFE), indicative of their intrinsic ability to form secondary structures under low water activity.

The XhLEA proteins did not show an effective role in protein stabilization when citrate synthase was dried in the presence of either of the proteins. It is likely that these proteins are involved in membrane stabilization through formation of a hydrogen bonding network

and/or a glassy matrix that could prevent membrane adhesion and stabilize the cytoplasm in the plasma membrane-cell wall continuum in the dry state. The hydrophilic nature of these XhLEA proteins, their expression in response to desiccation and the fact that they are plasma membrane/cell wall associated support the proposed function. The XhLEA proteins may provide a similar membrane stabilization role to PSVs of seeds. This is the first report in which the expression, in-solution structure and subcellular localization of *XhLEA1-4S1* and *XhLEA1-1S2* proteins in leaves, roots and seeds of *X. humilis* are described.

University of Cape Town

Table of Contents

DECLARATION.....	I
ACKNOWLEDGEMENTS	II
ABBREVIATIONS	III
FIGURES AND TABLES	IV
ABSTRACT.....	VIII
CHAPTER 1.....	1
<i>LEA PROTEINS AND DESICCATION TOLERANCE.....</i>	<i>1</i>
1.1 Introduction	1
1.2 LEA proteins.....	2
1.3 This study.....	21
CHAPTER 2.....	23
<i>PRODUCTION OF ANTIBODIES FOR XhLEA1-4S1 AND XhLEA1-1S2 AND EXPRESSION</i>	
<i>STUDIES.....</i>	<i>23</i>
2.1 Introduction	23
2.2 Methods and materials.....	25
2.2.1 Sub-cloning and expression of <i>XhLEA1-4S1</i> and <i>XhLEA1-1S2</i> in <i>E. coli</i>	25
A. Analysis of predicted amino acid sequence of <i>XhLEA1-4S1</i> and <i>XhLEA1-1S2</i>	
.....	25
B. Construction of expression plasmids	25
C. Transformation of expression host cells	26
D. Production of XhLEA recombinant proteins.....	26
E. Purification of XhLEA recombinant fusion proteins	27
F. Cleavage of GST tail off the recombinant LEA proteins	28
G. Confirmation of fusion protein identities	28
2.2.2 Production of polyclonal antibodies	29
2.2.3 Determination of XhLEA1 and <i>XhLEA1-1S2</i> expression conditions	30
B. Determination of RWC and protein extraction	30
C. Plant protein extraction	31
D. Western blot expression analysis	31

2.3 Results.....	32
2.3.1 Amino acid sequence analysis	32
2.3.2 Sub-cloning of <i>XhLEA1-4S1</i> and <i>XhLEA1-1S2</i> cDNAs into pGEX-3X expression vector.....	35
2.3.3 Expression of <i>XhLEA1-4S1</i> +GST and <i>XhLEA1-1S2</i> +GST recombinant fusion proteins	36
2.3.4 Production and purification of Polyclonal antibodies	42
2.3.5 <i>XhLEA1-4S1</i> and <i>XhLEA1-1S2</i> expression studies	44
2.4 Discussion.....	47
2.4.1 Sub-cloning of <i>XhLEA1-4S1</i> and <i>XhLEA1-1S2</i> cDNAs and production of recombinant proteins	47
2.4.2 Production and purification of polyclonal antibodies	48
2.4.3 <i>In planta</i> expression studies	50
CHAPTER 3.....	53
<i>IMMUNOLOCALIZATION OF XhLEA1-4S1 AND XhLEA1-1S2 PROTEINS IN X. HUMILIS</i>	
<i>LEAVES, ROOTS AND SEEDS.....</i>	53
3.1 Introduction	53
3.2 Materials and Methods	56
3.2.1 Tissue fixation	56
3.2.2 Immunogold labelling and electron microscopy	57
3.2.3 Counting of immunogold labels.....	57
3.2.4 Analysis of immunogold labelling data	58
3.3 Results.....	59
3.3.1 Analysis of gold labelling by Contingency Table	68
3.3.2 Analysis of gold labelling by Pair Correlation Function (PCF).....	75
3.4 Discussion.....	78
CHAPTER 4.....	81
<i>FUNCTIONAL AND STRUCTURAL CHARACTERIZATION OF XhLEA1-4S1 AND XhLEA1-1S2</i>	
<i>.....</i>	81
4.1 Introduction	81
4.2 Methods and materials.....	84
4.2.1 Construction of pET-21a(+) expression plasmids	84

4.2.2	Production of recombinant proteins	84
4.2.3	Fractionation of heat stable proteins	85
4.2.4	Isolation of T7-Tagged XhLEA recombinant proteins	85
4.2.5	Structural studies on <i>XhLEA1-4S1</i> and <i>XhLEA1-1S2</i>	86
4.2.6	LEA protein anti-aggregation role	87
4.2.7	LEA protein enzyme stabilization role	88
4.3	Results	88
4.3.2	Structural studies	92
4.3.3	Anti aggregation role of <i>XhLEA1-4S1</i> and <i>XhLEA1-S2</i>	97
4.3.4	Enzyme stabilization role of <i>XhLEA1-4S1</i> and <i>XhLEA1-S2</i>	99
4.4	Discussion	101
CHAPTER 5	106
GENERAL CONCLUSIONS	106
REFERENCES	109
APPENDICES	131

Chapter 1

LEA proteins and desiccation tolerance

1.1 Introduction

Green plants which are believed to have originated from a fresh water body before colonizing the terrestrial environment (Mishler and Churchill, 1985; Rensing *et al.*, 2008) had to develop multiple survival strategies over the ages to limit water loss and survive on a dry land. Internalizing phloem and xylem tissues, development of cuticular surface covering on leaves and stems and reducing leaf surface area are some of the structural adaptations used by higher plants to avoid excess water loss from vegetative tissues. Plants can also reduce water loss by controlling the rate of physiological processes. However, most plants cannot survive vegetative desiccation, instead, they produce desiccation tolerant seeds and spores that can germinate once favourable conditions are established.

Only few plants known as resurrection plants can survive vegetative desiccation to near dry state and fully recover on rehydration. This property was also reported in a few other anhydrobiotic organisms as described later in this chapter. Resurrection plants are subdivided into homoiochlorophyllous (HDT) and poikilochlorophyllous (PDT) groups. Whereas HDT retain their chlorophyll on desiccation, PDT plants lose their chlorophyll. Desiccation tolerance is a rare phenomenon in vascular and flowering plants but less exception in bryophytes. The full physiological recovery from seemingly dry and dead state on watering cannot be explained by the adaptations described above and is likely to involve multiple strategies.

Recent studies have shown, among others, up-regulation of messenger RNA (mRNA) transcripts of Late Embryogenesis Abundant (LEA) genes in response to desiccation and other abiotic stresses (Cuming *et al.*, 2007, Illing *et al.*, 2005; Collett *et al.*, 2004; Olave-

Concha *et al.*, 2004; Rabbani *et al.*, 2003). Therefore, LEA proteins might play a role in desiccation tolerance, however, little is known about these proteins. This is due to the fact that these proteins do not have structure under physiological conditions and hence are difficult to study. The available literature on LEA proteins, their classification, function and structural studies are summarised in this chapter.

1.2 LEA proteins

As the name implies, LEA proteins were originally discovered up-regulated in the late stages of orthodox seed development (Dure *et al.*, 1981; Galau and Dure, 1981; Galau *et al.*, 1986) and are a very broad grouping of proteins which can be grouped into different classes on the basis of sequence variability (described below in section 1.2.1). LEA proteins comprise about 4 % of cellular proteins during seed development (Roberts *et al.*, 1993). The common feature of LEAs is that they are low molecular weight (10 -30 kDa), low complexity, water deficit inducible proteins, unfolded in the hydrated state, extremely hydrophilic (with the exception of the D95 LEA family, Galau *et al.*, 1993) and heat stable; have no catalytic activity or structural domains; and most of them lack cysteine and tryptophan residues (Baker *et al.*, 1988; Lin *et al.*, 1990; Ingram and Bartels, 1996; Curry and Walker-Simmons, 1993; Close 1996; Bray, 1997). LEA proteins are boiling-soluble indicating that the proteins are hydrated and non-globular. These were the characteristics that led to the suggestion that LEA proteins are involved in the protection of plant cells from dehydration-induced damage (Zhang *et al.*, 2000).

The expression of LEA proteins in seeds is induced by the plant stress hormone, abscisic acid (ABA) and coincides with the onset of desiccation tolerance (Finkelstein *et al.*, 2002). In subsequent studies LEAs responsive to different stresses were reported from vegetative tissues of plants. Examples include LEAs from *Craterostigma plantagineum*, e.g. *CDT-1*, responsive to dehydration of vegetative tissues (Bockel *et al.*, 1998; Bartels *et al.*, 1990;

Bartels and Salamini, 2001), *COR15a* from *Arabidopsis thaliana* responsive to low temperature (Thomashow, 1999), *PpDHNA*, from *Physcomitrella patens* responsive to osmotic (salt) stress (Saavedra *et al.*, 2006) and *OSLEA3* from rice responsive to the application of exogenous ABA (Moons *et al.*, 1997).

Expression of LEA genes during seed maturation was reported in many plant species including barley, maize, rice, wheat, *Arabidopsis*, soybean, sunflower, potato, grape, apple, bean, tomato and rye (Baker *et al.*, 1988; Raynal *et al.*, 1989; Williams and Tsang, 1991; Almoguera and Jordano, 1992; Espelund *et al.*, 1992; Galau *et al.*, 1992; Litts *et al.*, 1992; Gaubier *et al.*, 1993; He and Fu, 1996; Yu, 2003; Han and Kermode, 1996; Wang *et al.*, 2003; Lan *et al.*, 2005; Illing *et al.*, 2005). Studies conducted on the recalcitrant seeds of *Avicennia marina* (Forsk.) Vierh., however, demonstrated that LEA proteins were not expressed in dehydration sensitive seeds (Farrant *et al.*, 1992) indicating that their expression in orthodox seeds might be associated with the acquisition of desiccation tolerance.

Illing *et al.* (2005) have analyzed the expression of 35 LEA genes from *Arabidopsis thaliana* during seed development and found out that 14 of them were seed specific. Of these, Group 1 LEAs were expressed during seed development only and not in response to other abiotic stress. Furthermore, one of these LEAs, a homologue of Group 6 LEA, was also expressed in dehydrated leaves of *X. humilis*, and it was suggested that such expression of desiccation associated LEAs might be as the result of the activation of seed specific genes.

It has been reported that LEA protein expression in seeds continues into the early stages of germination from stored mRNA of wheat embryos (Williamson and Quatrano, 1988; Morris *et al.*, 1990) and discontinues after few hours of germination (Roberts *et al.*, 1993). Others also reported up-regulated expression of LEA genes and down regulation of other genes

such as those involved in the cell cycle upon exposure to desiccation (Leprince and Buitink, 2010). This could be an indication that the expression of these desiccation linked LEAs is developmentally regulated in seeds and environmentally regulated in vegetative tissues.

There are many LEA proteins belonging to different groups and are broadly linked by their common expression in response to abiotic stresses or during seed development (Illing *et al.*, 2005; Dalal *et al.*, 2009). This might mean that either there are multipurpose LEAs or they have a subordinate role in stress tolerance, may be stabilizing other important macromolecules or enhancing the protective role of other molecules as previously suggested (Goyal *et al.*, 2005; Tunnacliffe and Wise 2007; Tompa and Kovacs, 2010).

Whether individual LEAs have specific function or form part of a group response against specific stress has not been sufficiently described. Manfre *et al.* (2009) demonstrated elevated expression of *A. thaliana* Group 1 LEA protein, *ATEM1*, in seeds of mutant *Arabidopsis* plants that did not express *ATEM6*, another Group 1 LEA protein. This might indicate common function or, as suggested by those authors, could indicate accelerated maturation program in the absence of the latter LEA. In the same report, the authors have indicated that *ATEM6* protein was not required in mature seeds for viability or efficient germination. It is possible that desiccation tolerance involves up-regulation of a number of genes and not just one type of LEA protein. Therefore it is important to investigate individual LEAs for different functions to determine their precise role in desiccation tolerance.

1.2.1 Classification of LEA proteins

The term LEA was assigned to these proteins originally to reflect their abundance in mature seeds, and had nothing to do with sequence similarity. Despite their common expression patterns and typical hydrophilicity (with the exception of the D95 LEA family, Galau *et al.*,

1993), LEA proteins are heterogeneous group and can be classified into groups on the basis of sequence similarity. The first classification attempt used the names of cotton, *Gossypium hirsutum* (Dure *et al.*, 1989). Subsequently, different authors have used different naming and umbering systems; for example Group 1 – 6 by Bray (1994) and Classes I – VI by Wise (2003). In this section, the present status of LEA protein classification is reflected and reference is made accordingly.

Based on the similarity of motif sequences a consensus has been established first for the three LEA protein groups namely Group 1 (D19), Group 2 (also known as dehydrins, D11) and Group 3 (D7) (Dure *et al.*, 1989). Three other groups including Group 4 (D113), Group 5 (D29) and Group 6 (D34) were added to the above list later, bringing the number of LEA protein grouping to 6 (Bray, 1994). However, those classification schemes were not all inclusive, and as shown in Wise (2003) (as Classes V, VI and III) and Illing *et al.* (2005) (LEA-7, LEA-8 and LEA-10), there are at least three other groups that were reported to be significantly different to form their own groups. This shows that the present status of LEA protein classification needs to be refined and consolidated.

Different classification approaches have been used by researchers in classifying LEA proteins into groups. For example; Garay-Arroyo *et al.* (2000) claimed that data base searching of LEA proteins using high glycine content (> 6 %) and a high hydrophilicity index (> 1.0) selectively identified most of the known LEA proteins. But Wise and Tunnacliffe (2004) described the search engines used as inadequate for LEA proteins due to the presence of low complexity sequences. The latter authors argued that the low complexity sequences of LEA proteins are masked to avoid skewing of statistical data attuned to globular proteins. Hence they preferred using a new computational method called POPP (Protein or Oligonucleotide Probability Profile; Wise, 2002 and 2003).

POPP is a software tool that allows the discovery of statistically unusual pattern in the composition of proteins for the purpose of clustering proteins into families based on peptide composition as opposed to similarities in their sequences. Based on this approach if charged and/or polar residues feature strongly then the protein is hydrophilic and vice versa (Wise and Tunnacliffe 2004). Using POPP, these authores confirmed the absence of cysteine residues in LEA proteins and also reported that phenylalanine, tryptophan, isoleucine, leucine and asparagine are under-represented. They also claimed that POPP clustering identified closely related proteins which resulted in reduced number of groups and also identified relatedness that could provide evolutionary significance.

In other studies *InterPro* / *Pfam* database search tools were used to classify LEA proteins into groups or families (Illing *et al.*, 2005; Fisher 2008; Battaglia *et al.*, 2008). In a classification presented by Battaglia *et al.* (2008), which was similar to previous classification by Dure *et al.* (1989), hydrophilic LEA proteins were categorized into seven families as Groups 1, 2, 3, 4, 6, and 7; and all the hydrophobic LEAs as Group 5. They indicated that their classification scheme allowed the identification of different families and motifs of LEAs conserved across species. Illing *et al.* (2005) and Fisher (2008) also classified LEA proteins *via InterPro* and *Pfam* scans. A summary table by Fisher (2008) showing how some of the different classification schemes relate to each other is given in Table 1.1 below.

Table 1.1 Nomenclature of LEA groups (Harada *et al.*, 1989; Dure, 1993; Galau *et al.*, 1993; Bray, 1994; Wise, 2003; Illing *et al.*, 2005) and their corresponding Pfam and InterPro descriptions and identifiers. *InterPro* descriptions are identical to *Pfam* descriptions unless otherwise noted in parentheses (from Fisher, 2008)

Dure 1993	Bray 1994	Wise 2003	Galau <i>et al.</i> , 1993	Harada <i>et al.</i> , 1989	Illing <i>et al.</i> , 2005	Pfam	InterPro
D19	Group 1	Class I	—	—	LEA-1	LEA_5 ; Small Hydrophilic Plant Seed Protein; PF00477	IPR000389
D11	Group 2	Class II	—	—	LEA-2	Dehydrin ; PF00257	IPR000167
D7	Group 3	Class III	—	—	LEA-3	LEA_4 ; PF02987 (LEA)	IPR004238
D113	Group 4	Class II, III	—	—	LEA-4	LEA_1 ; PF03760	IPR005513
D29	Group 5	Class III	—	—	—	LEA_4 ; PF02987 (LEA)	IPR004238
D34	Group 6	Class IV	—	—	LEA-6	Seed Maturation Protein ; PF04927	IPR007011
D73	—	Class V	Lea5	—	LEA-7	LEA_3 ; PF03242	IPR004926
D95	—	Class VI	Lea14	—	LEA-8	LEA_2 ; PF03168	IPR004864
—	—	Class III	—	Lea76	—	LEA_4 ; PF02987 (LEA)	IPR004238
—	—	—	—	—	LEA-10	AWPM-19-like ; PF05512	IPR008390

Despite the many attempts made, classification of LEA proteins on the basis of sequence similarity still remains confusing. An alternative approach would be to classify LEA proteins according to function. Function based classification would be of much more use and newly discovered LEAs would be investigated with a set of parameters on the basis of function. However, this is currently not possible due to the lack of sufficient functional studies reported in the literature.

In this study, the classification scheme proposed by Bray (1994), which was also used by Illing *et al.* (2005) and Fisher (2008), has been adopted (Table 1.1). Therefore, unless indicated otherwise, identification of LEA protein in this document refers to the Bray (1994) classification. In this latter approach, LEA proteins were classified into 6 groups, Group 1 – 6. Furthermore, the commonly studied LEA proteins belong to one of these groups. Specific characteristics of each of these groups are briefly discussed below.

- **Group 1 LEAs**

This group, also known as D-19 group, includes highly conserved and hydrophilic group of LEAs characterized by 20 amino acid motif sequence (Galau *et al.*, 1992; Close, 1997). This motif (GGQTRREQLGEEGYMQMGRK) has been shown to appear up to 4 times in the polypeptide (Close, 1996) and is believed to be the result of gene duplication followed by recombination or deletion (Tunnacliffe and Wise 2007).

The *EM* (for Early-methionine-labeled) protein, discovered in germinating wheat embryos (Cumming and Lane, 1979) belongs to Group 1 LEAs. Genes of this group are reported to have the highest degree of similarity at both nucleotide and amino acid sequence levels (Dure *et al.*, 1989). However, an *EM-like* LEA protein, (*OsLEA1a*), from rice (*Oryza sativa*) that doesn't have the 20-mer characteristic motif of Group 1 LEA proteins has been reported recently (Shih *et al.*, 2010c). Group 1 LEAs have been considered as plant specific; however, similar proteins have been reported from bacteria (*Bacillus subtilis*; Stacy and Aalen, 1998). A summary of non-plant proteins with evidence of expression is presented at the end of this section (Table 1.2).

In POPP approach Group 1 proteins are further divided into two Groups (Superfamilies: SF4 and SF6). Group1a (SF4) share POPP vectors with histone H4 and with other chromosomal and nuclear proteins with a predicted function of DNA binding, whereas Group 1b (SF6) match with a group of proteins with broader functions including RNA or ATP binding and gyrase or chaperone activities. Glutamate, arginine and Lysine residues are strongly represented in Group 1 LEA proteins (Wise and Tunnacliffe, 2004; Tunnacliffe and Wise, 2007).

- **Group 2 LEAs**

These group are commonly known as dehydrins or D-11 family (Dure 1993), also hydrophilic, highly unstructured and found in a variety of photosynthetic organisms including higher and lower plants, algae, and cyanobacteria (Close and Lambers, 1993; Close, 1997; Tompa and Kovacs, 2010). Group 2 LEAs are rich in Lysine and contain higher proportions of glutamate (Wise and Tunnacliffe 2004).

Group 2 LEAs are identified by the presence of at least three distinctive sequence motifs named as Y, S, K (Close *et al.*, 1993; Close, 1996; 1997). The K motif is made up of 15 amino acids (EKKGIMDKIKEKLPG) and found in the carboxyl terminus region, present up to 11 copies in a single polypeptide; whereas the shorter Y motif is found in the N terminus repeated up to 35 times. The S-segment is rich in Serine residues and is found phosphorylated in some proteins (Jiang and Wang, 2004). Based on the presence and arrangement of the Y, S and K motif, Group 2 LEAs are classified into 5 subgroups. K-subgroup are those that contain K-segment only and those that contain both S and K segments are in sub-group SK. Further variation in the S and K arrangements gave rise to YSK, YK and KS-subgroups (Campbell and Close, 1997).

The POPOs approach splits Group 2 LEAs into further grouping as Group 2a and Group 2b. Group 2a includes all Group 2 LEAs that are not associated with cold tolerance and those that are present late in embryogenesis; while Group 2b includes those that are generally associated with cold stress and those that are not produced during late embryogenesis. Group 2a overall charge is indicated as neutral or basic, with an over-representation of glycine. Group 2b also have similar levels of basic residues to Group 2a but with increased levels of acidic residues (Tunnacliffe and Wise, 2007).

- **Group 3 LEAs**

Group 3 LEA proteins are those that are identified by Dure (1993) as D7 and have 11-mer sequence motif (TAQAAKEKAGE) repeated upto 13 times in the polypeptide (also Bray 1993; Dure 2001) with high over-representation of Lysine and glutamate residues. This motif is recognized by apolar residues in positions 1, 2, 5 and 9, and charged or amide residues in positions 3, 6, 7, 8 and 11 (Dure 1993). Group 3 proteins are also unfolded in hydrated state; however, structure-prediction programs predict high degree of folding. In POPP based classification (Tunnacliffe and Wise, 2007), Group 3 LEAs include superfamilies 2 & 5 with some difference in their POPP consensus sequence. Examples include Barley PMA1949, carrot Dc8, and soybean pGmPM2 proteins (Yang *et al.*, 2002; Chen *et al.*, 2002; Sun, *et al.* 2003).

Group 3 LEA proteins are widely distributed in the plant kingdom and reported in some non-plant anhydrobiotic organisms such as bdelloid Rotifer *Philodina roseola* (*ProLEA1*; Tunnacliffe *et al.*, 2005), in the Arthropod *Polypedilum vanderplanki* (*PvLEA1*, *PvLEA2*, and *PvLEA3*; Kikawada *et al.*, 2006), in the insect *Megaphorura arctica* (Oac09516; Bahrndorff *et al.*, 2009), in the crustacean *Artemia franciscana* (*AfrLEA1* and *AfrLEA2*; Hand *et al.*, 2007) in the nematode *Aphelenchus avenae* (*AavLEA1*; Browne *et al.*, 2002; 2004). The expression of Group 3 LEAs in these organisms was correlated with desiccation tolerance.

- **Group 4 LEAs**

Group 4 LEA proteins were originally found highly accumulated in dry cotton embryos and classified as D113 by Dure (1993). They are identified by their conserved N-terminal motif, and are also disordered in aqueous solutions (Battaglia *et al.*, 2008). Transcripts of Group 4 LEA genes have been detected in soybean (Shih *et al.*, 2004), in wheat (Ali-Benali *et al.*,

2005) and in *Arabidopsis* (Delseny *et al.*, 2001; Bies-Etheve *et al.*, 2008; Dalal *et al.*, 2009). The N-terminal motif of this group is predicted to form amphipathic α -helical structure while the rest of the polypeptide remains unstructured (Campos *et al.*, 2006). As Group 3 LEAs, Group 4 LEAs are predicted to assume secondary structure in the presence of compounds such as 1% SDS or 50% TFE (Battaglia *et al.*, 2008). Other examples of Group 4 LEAs include the Cotton *D113*, *Craterostigma* *PGC27-45*, soybean *GmPM16* and tomato *LE25* (Galau *et al.*, 1993; Cohen *et al.*, 1991; Shih *et al.*, 2004).

- **Groups 5 and 6 LEAs**

The literature available on groups identified by Bray (1994) as Group 5 and 6 LEAs, or as D29 and D34 by Dure (1993) respectively as well as the additional groupings of LEAs seen in the other classification schemes (Table 1.2) is limited compared to the first 4 groups. Using Bioinformatics tools, Wise (2003) claims that Group 5 LEAs are misclassified and share more similarity with Group 3 (Class III; Table 1.2). The two also share the same *Pfam* and *InterPro* domains, *PF02987* and *IPR004238* respectively. In spite of this Group 5 remains a separate class of hydrophobic LEA proteins that adopt globular conformation (Battaglia *et al.*, 2008). *LE29_GOSHI* and *Q93Y63* are identified as members of this group by Bray (1993).

Group 6 LEA proteins on the other hand, are identified as *D-34* by Dure (1993) and include *PvLEA18* from bean (*Phaseolus vulgaris*) and found to be highly expressed at both transcript and protein levels in seeds, pollen grains and in response to dehydration and ABA treatment (Colmenero-Flores *et al.*, 1997). Another member of this group is LEA-6, a seed maturation protein from *A. thaliana*, whose expression was also found up-regulated in dry leaves of *X. humilis* (Illing *et al.*, 2005). The *Arabidopsis* genome is reported to contain six genes encoding *D-34* type proteins (Shih *et al.*, 2008).

Table 1.2 Non-plant LEA proteins with expression evidence (From Tunnacliffe and Wise, 2007).

LEA group	ID	AC	Seq Len	Expr	Species
Prokaryotes					
1	GSIB_BACSU	P26907	122	?	<i>Bacillus subtilis</i>
3	Q3EHT1_ACTSC	Q3EHT1	173	?	<i>Actinobacillus succinogenes</i> 130Z
3	UB72_DEIRA	Q9RV58	298	D	<i>Deinococcus radiodurans</i>
3	Q3XWV0_ENTFC	Q3XWV0	200	?	<i>Enterococcus faecium</i> DO
3	Q39ZB1_GEOMG	Q39ZB1	94	?	<i>Geobacter metallireducens</i> (strain GS-15)
3	Y1339_HAEIN	P71378	129	?	<i>Haemophilus influenzae</i>
3	Q11Z91_POLSJ	Q4AXW6	195	?	<i>Polaromonas</i> sp. JS666
3	Q2LPL8_SYNAS	Q2LPL8	113	?	<i>Syntrophus aciditrophicus</i> (strain SB)
Eukaryotes					
3	LEA1_APHAV	Q95V77	143	D	<i>Aphelenchus avenae</i>
3	AfrLEA1	*	357	D	<i>Artemia franciscana</i>
3	AfrLEA2	*	364	D	<i>Artemia franciscana</i>
3	Q61VH9_CAEBR	Q61VH9	775	?	<i>Caenorhabditis briggsae</i>
3	O16527_CAEEL	O16527	733	D	<i>Caenorhabditis elegans</i>
3	Q19790_CAEEL	Q19790	780	?	<i>Caenorhabditis elegans</i>
3	Q54VG7_DICDI	Q54VG7	319	?	<i>Dictyostelium discoideum</i>
3	Q8SVY9_ENCCU	Q8SVY9	166	?	<i>Encephalitozoon cuniculi</i>
3	Q1XI26_9DIPT	Q1XI26	742	D	<i>Polypedilum vanderplanki</i>
3	Q1XI25_9DIPT	Q1XI25	180	D	<i>Polypedilum vanderplanki</i>
3	Q1XI24_9DIPT	Q1XI24	484	D	<i>Polypedilum vanderplanki</i>

ID denotes SwissProt identifier; AC, accession number; Seq Len, protein sequence length. The Expr column indicates whether expression has been tested under conditions of desiccation stress (D) or otherwise (?). The sequences for the two proteins denoted by asterisks (*) can be found in Hand *et al.* (2007).

1.2.2 Functional studies on LEAs

Until recently, the main source of evidence for the many proposed roles of LEA proteins has been from observation and detection of high mRNA levels from transcriptome studies (Cuming *et al.*, 2007, Illing *et al.*, 2005; Collett *et al.*, 2004; Olave-Concha *et al.*, 2004; Rabbani *et al.*, 2003. In recent years, some functional-biochemical assays have been developed and ample evidence has been gathered to support the role of LEA proteins in desiccation tolerance as well as in other abiotic stresses. However, mechanisms of desiccation tolerance vary and our knowledge of how LEAs are involved with one or the

other mechanism is limited. Approaches used to investigate LEA protein function include *in-vitro* protection assays with purified recombinant LEA proteins and production of transgenic organisms where these proteins are overexpressed and tested for the proposed functions.

However, only few LEA genes have been used in transformation experiments involving, for example, yeast, bacteria, rice or tobacco and therefore the putative protective role of LEA proteins using transgenic organisms is not well investigated. Examples of LEA genes that have been used in such transformation experiments include *HVA1* of barley in rice (Xu *et al.*, 1996; Babu *et al.*, 2004); *BhLEA1* and *BhLEA2* of *Boea hygrometrica* in tobacco (Liu *et al.*, 2009); *PMA80* and *PMA 1959* of wheat in rice (Cheng *et al.*, 2002); *TaLEA2* and *TaLEA3* (Yu *et al.*, 2005) and *pYES2-Em* of wheat in yeast (Swire-Clark and Marcotte 1999; for more information, refer review by Tompa and Kovacs, 2010).

Gilles *et al.* (2007) reported that the N-terminal domain of Group 1 LEA protein is important in the proper folding of other proteins during dehydration and that it is involved in the protection of enzymes. They have produced a recombinant form of the wheat Group 1 (*rEM*) in *E. coli* capable of protecting the enzyme lactate dehydrogenase (LDH) from the deleterious effects of drying. When the LDH enzyme was dried in the presence of *rEm*, its activity was retained.

Manfre *et al.* (2006) also showed the importance of the *ATEM6*, a Group 1 LEA from *Arabidopsis thaliana*, for normal seed development. They performed an insertional mutation that resulted in failure of *ATEM6* protein expression which in turn resulted in premature drying of the seeds. This indicates the importance of this LEA in the development of desiccation tolerance in these seeds. Vegetative tissues of transgenic rice expressing wheat

Group 1 LEA protein (*PMA1959*) also maintained cellular integrity when subjected to salt stress (Cheng *et al.*, 2002). Minimum electrolyte leakage was measured from leaves of transgenic rice plants grown under salt stress conditions compared to wild type plants.

Over-expression of a wheat LEA1 protein (*EM*) in *Saccharomyces cerevisiae* conferred osmotic tolerance in yeast (Swire-Clark and Marcotile, 1999). These researchers have shown that yeast cells harboring wheat LEA1 protein showed un-inhibited growth in osmoticum compared to a control where no LEA protein was expressed; and is predicted to provide similar protection during water stress in plants. Lan *et al.*, (2005) also showed the osmoprotective role of another Group 1 LEA, *PM11* from soybean (*Glycine max* L. Merr. cv. Bainong 6), by growing *E. coli* expressing recombinant form of this protein.

The experimental findings described above indicate that Group 1 LEA proteins play protective role against not only desiccation caused by loss of water but also against osmotic stress caused by increased salt concentrations. However, the exact mechanisms on how these protections were provided were not adequately described. The characteristic 20mer motif of the group has high water binding property (Bray 1993; Close, 1996; Cuming, 1999) which is an important feature for tissues under water-deficit conditions, but whether this motif of Group 1 LEAs can account for all the observed and predicted functions is not established yet.

Group 2 LEA proteins (dehydrins), with the 15-amino acid conserved motif (EKKGIMDKIKEKLPG) at the C-terminus (Close 1997), are also known to play an important role in plant stress tolerance. Saavedra *et al.*, (2006) have shown up-regulation of *PpDHNA*, a Group 2 LEA, from *Physcomitrella patens* in response to salt and osmotic stresses. They demonstrated the role of *PpDHNA* against these stresses in a knockout mutant

where growth was severely impaired upon treatment with abscisic acid, NaCl or mannitol. Pair-wise overexpression of Group 2 LEA genes (*RAB18* and *COR47* or *LTI29* and *LTI30*) in transgenic *Arabidopsis* plant was also shown to play a role against freezing stress in *Arabidopsis* (Puhakainen *et al.*, 2004). It was suggested that such protection against freezing stress was due to the ability of these proteins to protect membrane damage.

The K-segment, that forms amphiphilic α -helix of Group 2 LEAs, is proposed to interact with membranes, partially denatured proteins and have synergetic action with compatible solutes in stabilizing macromolecules and protoplasm (Close 1996; 1997). Zhang *et al.*, (2000) used yeast expression systems to investigate the effect of Group 2 LEA proteins on cellular metabolism under different water stresses and found out that tomato *le4* (also Group 2) play a role in protecting cytoplasm from desiccation damage. Koag *et al.* (2003 and 2009) also proposed membrane stabilization role for maize DHN1, Group 2 LEA by conducting experiments involving phospholipid membranes (Koag *et al.*, 2003) and deletion of the K-segment (Koag *et al.*, 2009). They have reported that the latter segment promotes formation of α -helical structure and that the α -helical content of the protein DHN1 increases when the K-segment binds to a lipid vesicle.

Ectopic expression of wheat dehydrin (*DHN-5*) in *Arabidopsis thaliana* improved high salinity and water-deficit stress (Brini *et al.*, 2007b). The seedlings of this plant survived high salt stress treatment; and 30 % of seeds produced from the transgenic plants germinating after high salt treatment compared to 3 % from wild type. It has been reported that the distribution of Group 2 LEA proteins between the nucleus and cytoplasm is controlled by phosphorylation of the serine stutter and that the removal of this sequence results in the loss of these proteins from the cytoplasm (Jensen *et al.*, 1998). However, no

biochemical assays have been developed to measure the effects of the lack of this protein in the tissues.

Group 3 LEA proteins are also reported to play an important role in stress tolerance. In an *in vitro* experiment, a mitochondrial Group 3 LEA protein (*PsLEAm*) from a pea (*Pisum sativum*) was shown to interact with and protect liposomes subjected to drying (Tolleter *et al.*, 2007). Other Group 3 LEA proteins (*BN115m* and *BNECP63*) from *Brassica napus* prevented irreversible heat and freeze-induced precipitation of heat/freeze sensitive protein when expressed fused with it in *E. coli* (Singh *et al.*, 2009). In the same study, *BNECP63* was shown to facilitate expression of membrane chlorophyll-binding protein of photosystem II light harvesting complex (*BNPsbS*) and of a peptide of the Hepatitis C viral polyprotein (*HepC-CP*) recombinant proteins in *E. coli*. Furthermore, ectopic expression of *COR15A* in *Arabidopsis* provided protection against freezing damage of chloroplasts in both intact and isolated leaves (Thalhammer *et al.*, 2010).

Group 3 LEA proteins found in larvae of an African chironomid were also found to be related to the ability of the larvae to withstand almost complete desiccation during which it entered a state of suspended animation called anhydrobiosis (Kikawada *et al.*, 2006). Another Group 3 LEA recombinant protein, *AavLEA1*, has been shown to protect aggregation and activity of citrate synthase *in vitro* after repeated drying cycles (Goyal *et al.*, 2005). Chakrabortee *et al.* (2007) also reported an *in vitro* anti aggregation role for *AavLEA1* protein on polyQ protein in a mammalian cell lines using a similar method described by Goyal *et al.* (2005).

Expression of desiccation-inducible Group 3 LEA protein (Oac09516) was detected in *Collembola* species using antibodies generated against another Group 3 protein (*ArLEA1A*)

and was reported to be related to the desiccation tolerance of the insects (Bahrndorff *et al.*, 2009). Inactivation of Group 3 LEA proteins in bacterium, *Deinococcus radiodurans*, reduced viability of desiccated cultures by 75% (Battista *et al.*, 2001). Wolters *et al.* (2001) suggested that Group 3 LEA proteins play a role together with sugars in the formation of glassy cytoplasm during dehydration through tight hydrogen bonding network conferring long-term stability. Similar suggestions were also made by Shimizu *et al.* (2010).

Other functional roles predicted for Group 3 LEA proteins include sequestration of ions that would otherwise cause oxidative damage (Alsheikh *et al.*, 2003) or act as replacement for lost water in filling space (Hoekstra *et al.*, 2001), or assist in renaturing of unfolded proteins (Bray, 1991). Furthermore, these proteins were predicted to function as hydration buffers to maintain subcellular integrity (Cuming, 1999; Garay-Arroyo *et al.*, 2000).

Constitutive expression of a Group 4 LEA protein from *Arabidopsis thaliana* (*AtLEA4-5*) was shown to increase drought tolerance in this plant (Olvera-Carrillo *et al.*, 2010). In a similar study *BnLEA4-1*, a Group 4 LEA from *Brassica napus* conferred salt and heat tolerance when over expressed in transgenic *Arabidopsis* plants and *E. coli* (Dalal *et al.*, 2009). LE25, a Group 4 LEA protein from *L. esculentum* has been shown to improve growth rate for yeast cells at high KCl concentrations and improved tolerance against freezing (Imai *et al.*, 1996). At least in two more studies, Group 4 LEA proteins are reported to form a subsidiary structure adaptive to conformational changes of other proteins and function in protecting membrane stability and integration during drying (Li *et al.*, 1998; Chaves *et al.*, 2003).

Accumulation of *PvLEA18*, a Group 6 LEA protein in radicle of a germinating embryo was found to be associated with a protective function (Colmenero-Flores *et al.*, 1999). Over

expression of ASR1, a Group 7 LEA protein in tomato (Goldgur *et al.*, 2007), in tobacco (Kalifa *et al.*, 2004b) and in *Arabidopsis* (Yang *et al.*, 2005) resulted in increased salt tolerance.

From the functional studies reviewed here, it is evident that LEA proteins are associated with one or the other protective mechanisms against desiccation or other abiotic stresses and it seems that some LEA proteins have multiple or overlapping functions. This might indicate that LEA proteins function in an integrated fashion or it might be a reflection of our inadequate knowledge of how LEA proteins act under desiccation stress. It is also probable that many LEAs are required to protect plants against desiccation injury and that desiccation tolerant plants are required to express a specific set of desiccation-specific LEA proteins for this purpose. This would imply that desiccation sensitive plants do not express the required combination of LEA or other protective molecules.

1.2.3 Structural studies on LEAs

One approach used to investigate LEA protein function is to relate their structure with proposed functions. However, most of the LEA proteins investigated so far are reported to be unstructured at least in the hydrated state (Russouw *et al.*, 1995; Lisse *et al.*, 1996; Soulages *et al.*, 2002, 2003; Goyal *et al.*, 2003, Tolleter *et al.*, 2007; Shih *et al.*, 2010a and 2010c) and therefore difficult to crystallize. Hence, to investigate LEA protein structure, other alternative techniques including Circular Dichroism (CD) spectroscopy, Nuclear Magnetic Resonance imaging (NMR) or Fourier Transfer Infrared (FTIR) spectroscopy are used (Shih *et al.*, 2010a and 2010c).

Only few intrinsically disordered proteins are reported to adopt ordered structure as they bind to biological targets such as other proteins (Longhi *et al.*, 2003), membranes (Davidson *et al.*, 1998), RNA (Tompa and Csermely, 2004) or DNA (Love *et al.*, 2004). Group 1 and

Group 4 LEA proteins are predicted to exist as random coils (Soulages *et al.*, 2002) and contribute to stability during dehydration by sharing their hydration shell of water or by acting as water replacements through their hydroxylated amino acids (Manfre *et al.*, 2006).

Mouillon *et.al.* (2006) conducted an extensive structural investigation on four intrinsically disordered dehydrins (Group 2 LEAs) from *Arabidopsis thaliana* that are also expressed under conditions of water-deficit related stresses. These proteins have highly conserved stretches of 7-17 residues that are repetitively scattered in their sequences, the K-, S-, Y- and lysine rich segments. They investigated the putative role of these segments in promoting structure using CD spectroscopy.

Based on CD results obtained, Mouillon *et al.*, (2006) described LEA protein structure as random coil in highly hydrated state and increased proportion of helices as water is lost through drying to a collapsed structure enriched in α -helix or β -sheet as desiccation proceeds further. The structural transition during the drying period was attributed to the ability of LEA proteins to act as hydration buffers (also Manfre *et al.*, 2006). Wolkers *et al.* (2001) also reported that a Group 3 LEA protein (D-7) from *Typha latifolia* pollen, unordered in solution, assumed α -helical confirmation in the presence of high concentration of sucrose and was predicted to be involved in the formation of sucrose glasses.

Some researchers, however, suggested that the disordered nature of LEA proteins enable them to interact fast and flexibly through weak and reversible binding and to have more than one function (Kovacs *et al.*, 2008; Tompa *et al.*, 2005). As suggested by Mouillon *et al.* (2006) for the dehydrin proteins they studied, it seems that LEA proteins do not require acquisition of fixed tertiary structure for their biological function. Instead they interact with the macromolecules of the cell in a way that promotes functional and structural stability; and

the simplicity and unstructured nature of these proteins promotes the flexibility required to provide that stabilizing role.

Studies conducted on Early Responsive to Dehydration (ERD) 10 protein (Group 2 LEA) also indicated that the unstructured nature of proteins provides more surface area for more hydration compared to proteins with well folded and defined structure (Tompkins *et al.*, 2006). The availability of these LEA proteins in vegetative tissues of desiccation tolerant organisms therefore might ensure continued presence of the bound water, and may be very low level of metabolism similar to that of mature orthodox seeds.

It seems that LEA proteins remain unstructured under physiological conditions for as long as a critical moisture level is reached. Such understanding is supported by the finding that two Group 3 LEA proteins isolated from *Typha latifolia* pollen and a nematode changed their conformation from unordered to ordered structures during drying (Walker *et al.*, 2001). The unstructured nature of desiccation-linked LEAs under normal conditions also suggests that most of these proteins are not actively involved in physiological activities under normal conditions, however, this might not be true for LEAs responsive to other abiotic stresses.

The literature reviewed shows that considerable knowledge has been gained on LEAs that are expressed during seed embryonic development (Gaubier *et al.*, 1993; Cumming, 1999; Manfre *et al.*, 2006). However, there has been little comparative knowledge on the characteristics and expression patterns of LEAs produced in other organs of desiccation tolerant plant's leaves and roots when they are exposed to desiccation stress. In one report, a Group 6 LEA transcript, normally expressed during seed development, was found up-regulated in *X. humilis* vegetative tissues during desiccation (Illing *et al.*, 2005), however, there was no protein expression data to suggest that this LEA was translated into protein.

The presence of the different classes of LEA proteins (including those with limited sequence similarity), in diverse life forms, from bacteria to animals and plants, suggest that these LEA proteins play an important role in providing protection to terrestrial life forms against desiccation and other abiotic stresses. However, functional studies need to be done to elucidate the precise role of each of these LEA proteins in stress tolerance and desiccation tolerance in particular. A systematic and universal method of LEA classification would also facilitate research in this area.

1.3 This study

Group 1 LEA proteins are well characterized as being seed-specific in desiccation sensitive plants. However, it is not known if they are important in conferring desiccation tolerance to vegetative tissues of resurrection plants. *X.humilis* (Baker) T. Durand & Schinz, a resurrection plant, is used in our laboratory to investigate the phenomenon of desiccation tolerance. It is a low-growing perennial grass and is an indigenous Southern African plant. Ngubane (2008), a colleague in our lab, previously identified three Group 1 LEAs (XhLEA1-1, XhLEA1-2 and XhLEA1-4) in *X. humilis* using degenerate oligonucleotide primers and cloning. A fourth Group 1 LEA (XHC000797) was identified from a cDNA library of the same plant and annotated as seed specific (Arthur Shen, unpublished). Two of these LEAs, namely XhLEA1-4 and XHC000797, re-named as *XhLEA1-4S1* and *XhLEA1-1S2* respectively, were used for this study. Whereas the number '1' following XhLEA indicates that these proteins belong to Group 1 LEAs, the next numbers ('1' and '4') indicate the number of times the Group 1 LEA motif appears in the polypeptide sequence. 'S' stands for 'small hydrophilic plant seed protein signature which is repeated once in *XhLEA1-4S1* and twice in *XhLEA1-1S2*.

Messenger RNA data obtained by Ngubane (2008) suggested up-regulated expression of *XhLEA1-4S1* and *XhLEA1-1S2* in leaves and roots of *X. humilis* during desiccation. However, no protein expression data was obtained to suggest that these mRNAs were made into proteins and were functionally important. The aims of this study, therefore, were to characterize the expression of *XhLEA1-4S1* and *XhLEA1-1S2* proteins in leaves and roots of *X. humilis* during a cycle of desiccation and rehydration, to investigate structure and subcellular localization and to test the function of the recombinant proteins of these XhLEAs.

The full length cDNAs for *XhLEA1-4S1* and *XhLEA1-1S2* (Appendix A) were sub-cloned and the corresponding recombinant proteins were expressed in *E. coli*. These recombinant proteins were used for the generation of polyclonal antibodies which were used for the determination of expression conditions (Chapter 2) and immunolocalization studies (Chapter 3). Protein expression and localization were investigated by western blotting and gold labelling respectively. The in-solution structure of *XhLEA1-4S1* and *XhLEA1-1S2* was investigated by Circular dichroism and compared with results obtained by bioinformatic analysis tools (PONDR); and function of these proteins in desiccation tolerance was investigated by aggregation assay using the desiccation sensitive citrate synthase enzyme (Chapter 4). A general discussion on the findings of this work together with recommendations for future study of *XhLEA1-4S1* and *XhLEA1-1S2* or other similar proteins presented as a final chapter (Chapter 5).

Chapter 2

Production of antibodies for XhLEA1-4S1 and XhLEA1-1S2 and expression studies

2.1 Introduction

Approaches used in the study of plant stress tolerance include transcriptomic and proteomic techniques where cDNA libraries are made from target tissues and screened for differentially expressed genes in response to a given stress. The expression of candidate genes is tested at the mRNA level by techniques such as Northern blotting and Reverse Transcription quantitative Polymerase Chain Reaction (RT-qPCR) and/or at the protein level by Western blotting using antibodies. In this study, polyclonal antibodies were raised against *XhLEA1-4S1* and *XhLEA1-1S2* recombinant proteins to be used in western blotting for the determination of expression conditions of these XhLEA proteins in *X. humilis*.

The mRNAs of *XhLEA1-4S1* and *XhLEA1-1S2* LEAs were previously found up-regulated in the leaves and roots of *X. humilis* in response to desiccation (Ngubane, 2008), however, this was not confirmed by protein expression or localization data. Changes in mRNA levels may not necessarily correspond to actual protein synthesis, and most mRNA transcripts identified through this approach may not correspond to actual protein expression (Futcher *et al.*, 1999; Gygi *et al.*, 1999). It was reported that 50% of the transcripts found to be differentially expressed by transcriptomic approach actually didn't match to that of protein data (Watson *et al.*, 2003). One of the aims of this PhD study was therefore to investigate whether the pattern of *XhLEA1-4S1* and *XhLEA1-1S2* protein expression in roots and leaves in *X. humilis* in response to desiccation and subsequent rehydration was similar or different to that of mRNA transcript abundance.

There were few reports on the use of antibodies in the study of Group 1 LEA proteins. Swire-Clark and Marcotte (1999) used monospecific rabbit anti-*Em* antibody, made against wheat *Em*, to demonstrate the putative non-embryonic osmoprotective function of Group 1 LEA proteins in *Saccharomyces cerevisiae*. The results obtained indicate that the wheat *EM*-LEA protein can function as an osmoprotective molecule in yeast. In a similar study, antisera of LEA 1 raised against *Eleusine coracana* (Gaertn) protein, cross-reacted with a 27 *kDa* protein produced as a result of water loss due to polyethylene glycol (PEG) treatment of horsegram (*Macrotyloma uniflorum* (Lam.) Verde. Cv. VZM1) seeds (Veeranagamallaiah *et al.*, 2010).

LEA-specific antibodies were also used in the detection and determination of expression conditions of Group 1 LEAs in encysted embryos of brine shrimp, *Artemia franciscana* (Warner *et al.*, 2010). It was described that those LEA specific antibodies were generated against commercially synthesized peptides of the *Artemia* LEAs that were identified as good antigens and for their lack of glycosylation. Using Western blotting, they were able to detect at least three Group 1 LEA proteins designated as LEA-1a (21 *kDa*), LEA-1b (19 *kDa*) and LEA-1c (15.5 *kDa*) in the embryo cysts of *Artemia franciscana*. Similarly, Boudet *et al.*, (2006) raised polyclonal antibodies against recombinant *MtPM25* and *MtEm6*, Group 1 LEAs, to investigate the expression of these proteins in *Medicago truncatula* seeds in response to desiccation during seed maturation and germination. It was established that *MtPM25* and *MtEm6* proteins improved desiccation tolerance of cotyledons and radicles significantly.

Antibodies, raised against a synthetic peptide of 15 amino acids representing a consensus sequence (TGEKKGIMDKIKEKLPGQH) of dehydrin proteins (Group 2 LEAs), were also used to study the accumulation of these proteins during maturation of soybean [*Glycine max*

(*L. Merr.*] seeds in response to drought stress. It was found out that the accumulation of these dehydrins was maximal at seed physiological maturity (Samarah *et al.* 2006). In the current study, however, the *XhLEA1-4S1* and *XhLEA1-1S2* proteins were produced as GST fusion proteins and the whole fusion protein was injected into rabbits to generate the required antibodies against these proteins.

2.2 Methods and materials

2.2.1 Sub-cloning and expression of *XhLEA1-4S1* and *XhLEA1-1S2* in *E. coli*

A. Analysis of predicted amino acid sequence of *XhLEA1-4S1* and *XhLEA1-1S2*

Predicted amino acid sequences of *XhLEA1-4S1* and *XhLEA1-1S2* were scanned for conserved domains and other patterns against PROSITE database available at ExPASy Proteomics Server (<http://ca.expasy.org/prosite/>). Multiple sequence alignment was done by COBALT sequence alignment tool, (<http://www.ncbi.nlm.nih.gov/tools>).

B. Construction of expression plasmids

The standard protocols from the GST Gene Fusion System manual (Amarsham Pharmacia Biotech, 1997) were used for the sub-cloning, in pGEX-3X, and expression of *XhLEA1-4S1* and *XhLEA1-1S2* in *E. coli* unless specified otherwise. The full-length nucleotide sequences of *XhLEA1-4S1* and *XHLEA1-1S2* obtained previously by PCR amplification from seed and root cDNAs respectively (Ngubane 2008) were used as templates for sub-cloning. Forward and reverse PCR primers were designed to amplify these full length cDNAs (Table 2.1). BamH1 and EcoR1 restriction sites were included in the primers to facilitate subcloning of the XhLEA PCR products into the pGEX-3X expression vector. (PCR: 30 cycles of 94 °C, 60 °C, 72 °C, and a final extension at 72 °C for 5min). PCR products were restriction digested, gel purified (QIAquick Gel Extraction Kit, Cat. No. 28704) and ligated into the target pGEX-3X expression vector (pGEX-3X vector map, Appendix B). (Note that *XhLEA1-4S1* and *XHLEA1-1S2* cDNAs were also cloned into pET-

21(a+) vector using this approach for the production of recombinant proteins for purposes of functional and structural studies on these proteins, cloning in pET -21(a+) and protein production are described in Chapter 4).

Table 2.1 Primers used for the amplification of the full-length *XhLEA1-4S1* and *XhLEA1-1S2* cDNAs for subcloning into pGEX-3X. BamH1 and EcoR1 restriction sites (underlined) were included to facilitate in frame cloning into the pGEX-3X.

Primer name	Sequence 5' --- 3'
<i>XhLEA1-4S1</i> Forward primer:	G <u>TAGGATCCT</u> TATGGCTTCCCATCAA
<i>XhLEA1-4S1</i> Reverse Primer:	GCTG <u>AATTCA</u> CGAACTAAGA ACGTCG
<i>XhLEA1-1S2</i> Forward Primer:	G <u>TAGGATCCT</u> TATCAATCATGGCTTCTG
<i>XhLEA1-1S2</i> Reverse Primer:	GCATG <u>AATTC</u> CATCTCGTCGGAGATAGGT

C. Transformation of expression host cells

Preparation and transformation of competent *DH5-alpha* (Invitrogen) and *BL21 (DE3)pLysS* (Novagen) *E.coli* cells were done according to the supplier's instructions. Ligation reaction product of either pGEX-3X+*XhLEA1-4S1*- or pGEX-3X+*XhLEA1-1S2* from procedure "2.2.1 B" above was used first to transform competent *DH5-alpha* cells reported to be easier to transform (Taylor *et al.*, 1993). Selected colonies were PCR-screened for the presence of the recombinant plasmids. Recombinant plasmids were then isolated from positive colonies, sequenced and used to transform the preferred *BL21(DE3)pLysS* protein expression host cells.

D. Production of XhLEA recombinant proteins

Starter cultures of 5 ml of 2 X YT medium (Yeast extract and Tryptone) containing 100 µg/ml ampicillin and 34 µg/ml chloramphenicol were inoculated each with a single *BL21(DE3)pLysS* positive colony containing the recombinant pGEX plasmid of either *XhLEA1-4S1* or *XhLEA1-1S2* and incubated overnight at 37 °C with vigorous shaking. This

culture was diluted to 1:100 in a total volume of 100 ml with fresh pre-warmed media and grown until the OD₆₀₀ reached 0.6. Protein expression was induced by the addition of isopropyl β -D-thiogalactoside (IPTG) to a final concentration of 1 mM. Optimum induction time for expression of recombinant proteins was determined by taking 1 ml aliquots of the cultures just before IPTG addition and at 30, 60, 90, 120, 180 and 240 minutes after induction.

Cells were pelleted by spinning at 7700 x g for 10 minutes and re-suspended in phosphate buffered saline (1 x PBS; 140 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄, pH 7.3), containing a protease inhibitor (PMSF, 1 mM final concentration), and lysed by sonication in an ice/ethanol bath with a microtip attached to a sonicator (Misonix 3000), output power set at 4, 5 cycles of alternating 30 seconds ON and 60 seconds OFF to decrease foaming. Lysates were spun at 12000 x g for 10 minutes and the supernatant was transferred to new tubes. Protein concentration was determined using Bradford method (Bradford, 1976) and 10 μ g of protein was loaded on 12 % SDS-PAGE for analysis. The conditions for optimal expression of recombinant proteins in 100 mls were applied to cultures of 200 mls for the large scale production of recombinant protein.

E. Purification of XhLEA recombinant fusion proteins

The supernatants of GST-*XhLEA1-4S1* and GST-*XhLEA1-1S2*, of 200 ml cultures each from step “C” were filtered through a 0.45 μ m millipore membrane and loaded onto separate Glutathione–Sepharose 4B columns (GST Purification Module, 27-4570-01) equilibrated with 1 x PBS at 4 °C. Unbound protein was washed through with 10 volumes of 1 x PBS and the GST-fusion protein was eluted according the manufacturers protocol with 20 mM reduced glutathione in 100 mM Tris–HCl, (pH 8.0). 5 serial elutions, each with 1 ml elution buffer, were performed. Elutes were pooled and protein enrichment and desalting was done

using 10 *kDa* MWCO columns (Microsep, Cat. No. OD010C41). Protein concentration was determined (Bradford method) and quality was analyzed on 12 % SDS–PAGE.

F. Cleavage of GST tail off the recombinant LEA proteins

Cleavage experiments were performed on both bound and eluted GST-LEA fusion proteins with Factor *Xa* (GE Healthcare, Cat. No. 27084901). For bound proteins, the fusion protein-bound matrix was washed with 10 bed volumes of Factor *Xa* Cleavage Buffer (50 µl of Factor *Xa* Solution with 950 µl of 1X PBS for each ml of glutathione Sepharose bed volume). The Factor *Xa* mixture was added to the bound fusion protein in the ratio of 1 % (w/w) factor *Xa* to substrate. The suspension was gently rotated at room temperature for 10 hours after which the eluate was collected. Similarly, 10 cleavage units of Factor *Xa* solution per mg of fusion protein was added to elute GST-LEA fusion protein, mixed gently and incubated as before and the cleaved GST was removed by column purification method whereby the GST remains bound in the column and freed LEA proteins eluted. Elutes from both cleavage experiments were analyzed by SDS-PAGE.

G. Confirmation of fusion protein identities

Identity confirmation of GST-XhLEA fusion proteins was done by Mass Spectrometry (MS) at the Centre for Proteomic and Genomic Research, University of Cape Town. Affinity purified fusion proteins were loaded on 12 % SDS gel and the specific bands were cut out, processed and trypsin digested. Digested peptides were spotted onto the MALDI source plate and analyzed with a 4800 MALDI ToF/ToF (Applied Biosystems). All MS spectra were recorded in positive reflector mode. Spectra were generated with 600 laser shots/spectrum at laser intensity of 3600 (arbitrary units) with a grid voltage of 16 kV. All peptide containing spots were internally calibrated using trypsin autolytic fragments. Experimental spectra were manually compared with peptides of theoretical digestion and the sequences corresponding to observed ions assigned

2.2.2 Production of polyclonal antibodies

A. Rabbit immunization and collection of antiserum

Affinity purified GST-*XhLEA1-4S1* and GST-*XhLEA1-1S2* fusion proteins were administered each to a separate New Zealand White rabbit according to the procedure adopted by the Animal Unit, University of Cape Town (Rybicki, 1979; see protocol outline in Appendix C). 10 ml pre-immune serum was collected prior to the administration of the antigen from the marginal ear vein. Rabbits were then inoculated subcutaneously on the back with either GST-*XhLEA1-4S1* or GST-*XhLEA1-1S2* proteins (antigens, 500 µg/injection) in the presence of Freund's incomplete adjuvant. (Antigen injection, monitoring of rabbits before and after injection, bleeding and euthanizing were done by UCT Animal Unit personnel).

The first injection was followed by 3 subsequent booster injections at weekly intervals for a total antigen of 2 mg per rabbit. The development of antibody was monitored by collecting blood every two weeks from the rabbits and checking for the presence of antibodies using standard ELISA (Enzyme-Linked Immunosorbent Assay). The end point titre was estimated as described previously (Bishop *et al.*, 1984; Classen *et al.*, 1987; Frey *et al.*, 1998) as the dilution above the cut-off point that gives positive result. The OD readings for each bleed were then corrected by deducting the ODs of the negative controls for each dilution and the cut-off point was calculated as the mean of negative controls plus 3SD (SD =standard deviation).

B. Polyclonal antibody purification

Antibodies were purified with a sequence of steps. First IgG antibodies were precipitated with saturated ammonium sulphate (Sigma, Antisera Protocols). Precipitated *IgG antibodies* was dissolved in phosphate buffer and the suspension was passed through DE52 column, pre-equilibrated with the same buffer, and 1 ml elute fractions were collected. Concentration

of elutes was measured (OD_{280} of 1.4 = 1 mg/ml) and samples with higher OD pooled. The partially purified IgG antibody solution was then dialysed against two changes of 1X PBS using a dialysis tubing (12 kDa MWCO, Sigma, Cat. # D9652). 1 ml of the dialysates of each serum type was first incubated with GST-bound glutathione beads to remove GST specific antibodies. Flow through from this step was then incubated with the specific GST-LEA fusion protein bound glutathione beads. After repeated washes, LEA specific antibodies were eluted with 0.2 M glycine, pH 2.5. Eluates were neutralized with 1M K_2HPO_4 and stored in aliquots of 50 μ l at -80°C.

2.2.3 Determination of XhLEA1 and XhLEA1-IS2 expression conditions

A. Plant material, dehydration and rehydration treatments

X. humilis plants were collected from Pilanesberg National Park and Borakalalo National Park, South Africa and maintained in a glasshouse at the University of Cape Town under ambient conditions as described in Dace *et al.* (1998) prior to experimentation. Fully hydrated plants were then moved to a phytotron set with the following growth conditions: Temp. 24 °C, light/dark cycles of 14/10 hours, light intensity 188 μ mol m⁻²s⁻¹, humidity +/- 50 %], and allowed to acclimatize for 3 weeks. Drying was initiated by withholding water during which samples were collected for determination (see below). After reaching the air-dry state (< 10% RWC, reached in 12 days) plants were kept dry for 2 more weeks before rehydration by soil watering and sample leaves were collected. Control plants were kept fully hydrated by regular watering.

B. Determination of RWC and protein extraction

Absolute water content (AWC) of leaf samples was calculated using the formula: (fresh biomass–dry biomass)/dry biomass, which was then used to determine the Relative Water Content (RWC) using the formula: (AWC \times 100)/AWC at full turgor (Sherwin and Farrant 1996). AWC of samples watered the previous night and measured early in the morning

before temperature-induced transpiration began was used as the AWC at full turgor. 15 leaves were randomly collected at each time point, five of which were used for the determination of water content and the remaining leaves were flash-frozen and kept at -80°C . The RWC of the remaining 10 leaves was assumed to be the average of the 5 leaves.

C. Plant protein extraction

Protein was extracted from the 10 leaves, collected as described in Step 2.2.3.B above, at the following RWCs during dehydration treatment: fully hydrated, 75 %, 55%, 35%, 20% and air dry (<10% RWC) as well as after 6, 18, and 42 hours of watering using Plant Hydrophilic Protein Extraction Kit (Sigma, Cat. No. PE0210) following the supplier's instructions. Protein was also extracted from fully hydrated roots; desiccated roots and seeds of *X. humilis*. A variety of protein extraction protocols were attempted and it was found out that the Hydrophilic Protein extraction kit gave the best yield for the purpose intended. Protein concentrations were determined using the Bradford (1976) method and integrity was analyzed on SDS-PAGE. These proteins were used for western blot expression studies. Protein extracted from *Arabidopsis thaliana* seeds (provided by Dr. Rob Ingle), extracted using the same method, was also included for comparison purposes.

D. Western blot expression analysis

20 μg protein, from extractions described in step 2.2.3A above, was loaded on duplicate 12 % SDS-PAGE and each transferred onto a PVDF membrane prepared according to the recommended manufacturer's protocol (Roche, Cat. No. 03010040001). Protein transfer was done using standard procedure (240 mA constant current for 1 hr 45 minutes). Each membrane was exposed to either anti-*XhLEA1-4S1* or anti-*XhLEA1-1S2* polyclonal antibodies, and at least three technical repeat western blot experiments were conducted. Membranes were blotted, developed and analyzed using SuperSignal West Pico Chemiluminescent substrate Kit (Thermo Scientific, Cat. No. 34083).

2.3 Results

2.3.1 Amino acid sequence analysis

When translated, *XhLEA1-4S1* encoded for a 16.5 kD protein with an iso-electric point (pI) of 6. *XhLEA1-1S2* translated into 13.9 kDa polypeptide of pI of 5.3. Both the proteins contain the characteristic small hydrophilic plant seed protein signature G-[EQ]-T-V-V-P-G-G-T (PROSITE Database, Entry PS00431) (sequence in box, Figure 2.2) appearing once in *XhLEA1-4S1* (from amino acid number 17 to 25) and twice in *XhLEA1-1S2* (17 to 25 and 91 to 99). This PROSITE motif is also present in other Group 1 LEA proteins from *X. humilis* and other angiosperms and moss and where it appears only ones (Appendix D). The characteristic 20 amino acid long motif (consensus sequence: GGETRREQLGSEGYSEIGHK) of Group 1 LEA proteins is also present, repeated 4 times in *XhLEA1-4S1* and once in *XhLEA1-1S2* (Figure 2.1).

Furthermore, *XhLEA1-4S1* protein contains Y[ETQ][ED][IMV] motif (amino acid number 60 – 63) which is also present in *A. thaliana EM1* and another stress related protein (Database code for the latter protein is Q949P1, UniProt). The latter protein was reported to be involved in determining abscisic acid levels in dry seeds and in the control of post germination growth (Saito *et al.*, 2004). This motif was also reported to have a role in specifying intracellular signalling (Songyang *et al.*, 1994). The latter motif is present in *A. thaliana EM1* protein, but not in *XhLEA1-1S2* protein.

A

<i>XhLEA1-4S1</i>	
1	TTGTGAGCAGGCATGCGCTTCCCATCAAGAAAGGGCTGAGATGGACCGTAGGGCCAGGGAAGGTGAGACTGTCGTACCCGGTGGTACCGGC
1	L * A G M A S H Q E R A E M D R R A R E <u>G E T V V P G G T</u> G
91	GGAAAGAGCCTCGAAGCTCAACAACACCTCGCCGAAGGTCGTAGCAAGGGAGGGCAGACGAGGAAGGAACAGTTGGGGACAGAGGGGTAT
31	G K S L E A Q Q H L A E G R S K <u>G G Q T R K E Q L G T E G Y</u>
181	CAGGAGATGGGCCACAAAGGTGGTGAGACGAGGAGGGAACAGCTGGGTTCAGAAGGGTACAGTGAATCGGGCACAAAGGAGGGGAGGCG
61	<u>Q E M G H K G G E T R R E Q L G S E G Y S E I G H K G G E A</u>
271	AGGAGAGAGCAACTGGGTCTGAAGGGTACAGTGAAATCGGGCACAAAGGTGGTGAGACGAGGAGGGATCAGATTGGGTTCGGAGGGATAC
91	<u>R R E Q L G S E G Y S E I G H K G G E T R R D Q I G S E G Y</u>
361	CGTGAGATGGGTTCGAAAGGTGGTCTGTGACCAAGGACGAGTCCGGGGGAGAAGCTGCTGCCCGGAGGGCATAGAGATCGACGAGTCC
121	<u>R E M G R K G G L S T K D E S G G E R A A R E G I E I D E S</u>
451	AAGTACAGGACCAATGTTTAAGGTCTTGCGACGTTCTTAGTTCGTGGTATTAGTAGTGTGTGTGTGTGTGTATCGTATGTGTGTTA
151	K Y R T N V * G L A T F L V R G I * * L C V V L C I V C V L

B

<i>XhLEA1-1S2</i>	
1	GGGAACATCAATCATGCGTTCTGCACAGGAGAGGATTGAGCTCGACCGAAGGGCGAGGGAGGGCGAGACTGTTGTCCCCGGCGGCACAGG
1	G T S I M A S A Q E R I E L D R R A R E <u>G E T V V P G G T</u> G
91	CGGAAAGAGCCTCGAAGCTCAAGAACACCTCGCTGACGGACGGAGCCGTGGAGGGCAGACGCGCAGGGATCAGCTGGGATCAGAAGGGTA
31	G K S L E A Q E H L A D G R S R <u>G G Q T R R D Q L G S E G Y</u>
181	CAGTGAGCTTGCCCGCATGGGTGGACAGAGCGGCGGCTCGTATTCCGATGAGACGCCACCGCGGGGTGTTGGGGCCCGATCTCGGAGC
61	<u>S E L G R M G G Q S G G S Y S D E T A T G G V L G P D L G A</u>
271	AGGGGAGACGGTTCGTACCTGGAGGCGCGGCGGGAAGAGTGTGAAGCTCAAGAGAATCTCGCCGGAGGACGGAGACCGGTGGGGAGAC
91	<u>G E T V V P G G A</u> G G K S V E A Q E N L A G G R R R G G E T
361	ACGCATAGAGCAGCTTGGTGCCGAAGGCTACAGTGCACTCGGCCGACGGGGGAAGCAGCTGAACGCCACCTATCTCCGACGAGATCGA
121	R I E Q L G A E G Y S A L G R T G G S S * T P P I S D E I

Figure 2.1 Predicted amino acid sequences of *XhLEA1-4S1* (A) and *XhLEA1-1S2* (B). The nucleotides included in the forward and reverse primers as well as the 20 amino acid repeating motif are underlined. Translation start codon is indicated in blue colour fonts. * is used for stop codons. The small hydrophilic plant seed protein signature is put in box.

The sequences of *XhLEA1-4S1* and *XhLEA1-1S2* showed high similarity at both nucleotide and amino acid levels (57% and 53 % respectively, E-value for protein identity = 4e-28). The predicted amino acid sequences of these XhLEA proteins were used for database identity and similarity search and showed an identity of 76% (E value = 3e-57) and 51% (E value = 1e-24) for *XhLEA1-4S1* and *XhLEA1-1S2* to *EM1* LEA protein from *Arabidopsis* (NP_190749.1, NCBI/ BLAST/ blastp suite) which also contains the 20 amino acid motif of Group 1 LEAs (Figure 2.2).

<i>EM1</i>	1	MASKQLSREELDEKAKQGETVVPGGTGGHSLEAQEHLEGRSKGGQTRKEQLGHEGYQEIGHKGGEA
<i>XhLEA1-4S1</i>	1	MASHQ-ERAEMDRRAREGETVVPGGTGGKSLEAQOHLAEGRSKGGQTRKEQLGTEGYQEMGHKGGET
<i>XhLEA1-1S2</i>	1	MASAQ-ERIELDRRAREGETVVPGGTGGKSLEAQEHLADGRSRGGQTRRDQLGSDGYSELGRMGQS
<i>EM1</i>	68	RKEQLGHE-----GYQEMGHKGGEARKEQLGHEGYQEMGHKGGEARKEQLGHEGYKEMGRKG
<i>XhLEA1-4S1</i>	67	RREQLGSE-----GYSEIGHKGGEARREQLGSEGYSEIGHKGGETRRDQIGSEGYREMGRKG
<i>XhLEA1-1S2</i>	67	GGSYSEMATGGVLPDLGAGETVVPGGAGGKSVEAQENLARGRRRGGETRREQLGTEGYLSALGRTG
<i>EM1</i>	125	GLSTMEKSGGERAE EEGIEIDESKF-TNK
<i>XhLEA1-4S1</i>	123	GLSTKDESGGERAA REGIEIDESKYRTNV
<i>XhLEA1-1S2</i>	134	GSS-----

Figure 2.2 Sequence alignments of *XhLEA1-4S1* and *XhLEA1-1S2* against *A. thaliana* EM1 protein. Shared amino acid or peptide sequences among the three proteins are shaded in blue, those that are shared by *EM1* and *XhLEA1-4S1* are shaded brown and between *XhLEA1-4S1* and *XhLEA1-1S2* are shaded in pink. *EM1* and *XhLEA1-4S1* share the highest sequence similarity.

Amino acid sequence analysis also revealed that cysteine, tryptophan or phenylalanine residues are absent in both *XhLEA1-4S1* and *XhLEA1-1S2* proteins as is the case for many known LEA proteins; and Glutamic acid, Glycine and Arginine residues are over represented (Houde *et al.*, 1992) making up about 45 % of the total weight of both polypeptides. Asparagine in both XhLEAs, Proline and Valine in *XhLEA1-4S1*; Histidine, Isoleucine and Lysine in *XhLEA1-1S2* are under represented (Table 2.2).

Table 2.2 Amino acid composition of *XhLEA1-4S1* and *XhLEA1-1S2* proteins.

<i>XhLEA1-4S1</i>					<i>XhLEA1-1S2</i>				
Amino acid		Num	%mol/mol	%w/w	Amino acid		Num	%mol/mol	%w/w
A	Ala	8	5.26	3.70	A	Ala	11	8.09	6.00
C	Cys	0	0.00	0.00	C	Cys	0	0.00	0.00
D	Asp	4	2.63	2.76	D	Asp	6	4.41	4.89
E	Glu	25	16.45	19.09	E	Glu	14	10.29	12.62
F	Phe	0	0.00	0.00	F	Phe	0	0.00	0.00
G	Gly	31	20.39	12.08	G	Gly	32	23.53	14.71
H	His	5	3.29	4.03	H	His	1	0.74	0.95
I	Ile	5	3.29	3.40	I	Ile	1	0.74	0.80
K	Lys	9	5.92	6.83	K	Lys	2	1.47	1.79
L	Leu	6	3.95	4.08	L	Leu	10	7.35	8.03
M	MET	4	2.63	3.10	M	MET	3	2.21	2.74
N	Asn	1	0.66	0.69	N	Asn	1	0.74	0.81
P	Pro	1	0.66	0.60	P	Pro	3	2.21	2.12
Q	Gln	9	5.92	6.83	Q	Gln	7	5.15	6.27
R	Arg	17	11.18	15.36	R	Arg	16	11.76	17.07
S	Ser	11	7.24	6.00	S	Ser	12	8.82	7.73
T	Thr	8	5.26	4.94	T	Thr	8	5.88	5.84
V	Val	3	1.97	1.82	V	Val	6	4.41	4.30
W	Trp	0	0.00	0.00	W	Trp	0	0.00	0.00
Y	Tyr	5	3.29	4.70	Y	Tyr	3	2.21	3.33

2.3.2 Sub-cloning of *XhLEA1-4S1* and *XhLEA1-1S2* cDNAs into pGEX-3X expression vector.

PCR with the primers designed for the sub-cloning of *XhLEA1-4S1* and *XhLEA1-1S2* cDNAs produced DNA bands with the expected sizes (504bp for *XhLEA1-4S1* and 456bp for *XhLEA1-1S2*, Figure 2.3). Removal of other contaminating DNA bands observed as smears and faint bands was required; and this was done by cutting out the bands of interest followed by purification. The PCR products were successfully ligated, each into similarly restriction digested pGEX-3X vector.

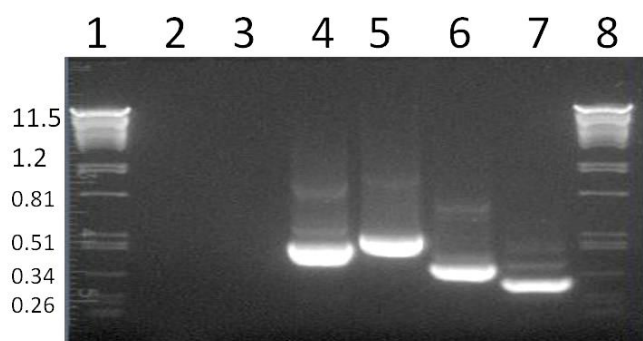


Figure 2.3 Insertion of *Bam*HI and *Eco*RI enzyme restriction sites into LEA cDNA clones through PCR. Lanes 1 and 8 = PstI-λ DNA marker, lanes 2 and 3 = no template and no primer controls respectively. lane 4 = *XhLEA1-1S2*, lane 5 = *XhLEA1-4S1*, lane 6 and 7 cDNAs of *XhLEA1-1* and *XhLEA1-2* (that were not used in subsequent work).

Transformation of *DH5-alpha* cells was successful. This was demonstrated by colony-PCR (Figure 2.4), from which the respective recombinant plasmids of *XhLEA1-4S1* and *XhLEA1-1S2* were isolated. The transformation of preferred expression host cells, *BL21(DE3)pLysS* cells, was also successful when checked as above. When selected colonies were checked, DNA bands of expected sizes were observed on 1 % agarose gel (Figure 2.5). Sequence results of *XhLEA1-4S1* and *XhLEA1-1S2* in pGEX-3X plasmids obtained before being used in the transformation of *BL21(DE3)pLysS* cells and after transformation did not show any mutation. Plasmids from clone 5 for *XhLEA1-4S1* and clone 9 for *XhLEA1-1S2* were sequenced (Figure 2.5).

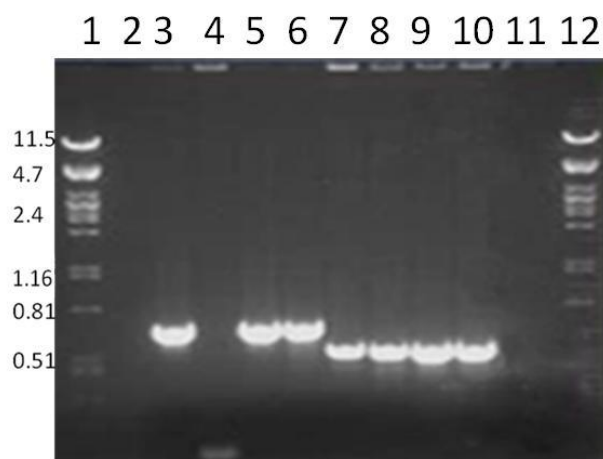


Figure 2.4 Colony screening by PCR for *pGEX-XhLEA1-4S1* (lanes 3-6) and *pGEX-XhLEA1-1S2* (lanes 7-10) recombinant plasmids in *DH5-alpha* cells using pGEX forward and reverse primers. Lanes 1 and 12 are Pst- λ DNA marker and lanes 2 and 11 are no-template and no-primer controls respectively.

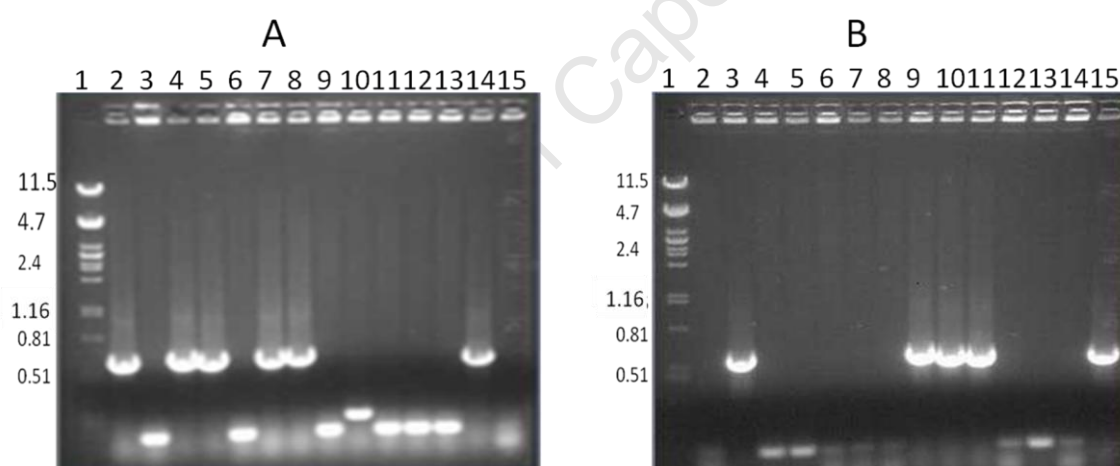


Figure 2.5 Colony screening for *pGEX-XhLEA1-4S1* (A) and *pGEX-XhLEA1-1S2* (B) recombinant plasmid in *BL21(DE3)pLysS* colonies. Lane 1 in both A and B represent Pst- λ DNA marker, Lanes 2 to 15 in both Figures represent number of colonies screened.

2.3.3 Expression of XhLEA1-4S1+GST and XhLEA1-1S2 +GST recombinant fusion proteins

A differential expression of both fusion proteins was observed after 30 minutes of IPTG addition and continued to increase in concentration for up to 4 hours tested (Figure 2.6). The size of the protein bands could not be measured accurately from SDS gels as the migration is known to be affected by a number of factors including molecular weight, however, from the

migration of the molecular weight marker, the bands representing the recombinant fusion proteins were located around the expected sizes (43 and 40.5 *kDas* for XhLEA1-4S1-GST and XhLEA1-1S2 respectively).

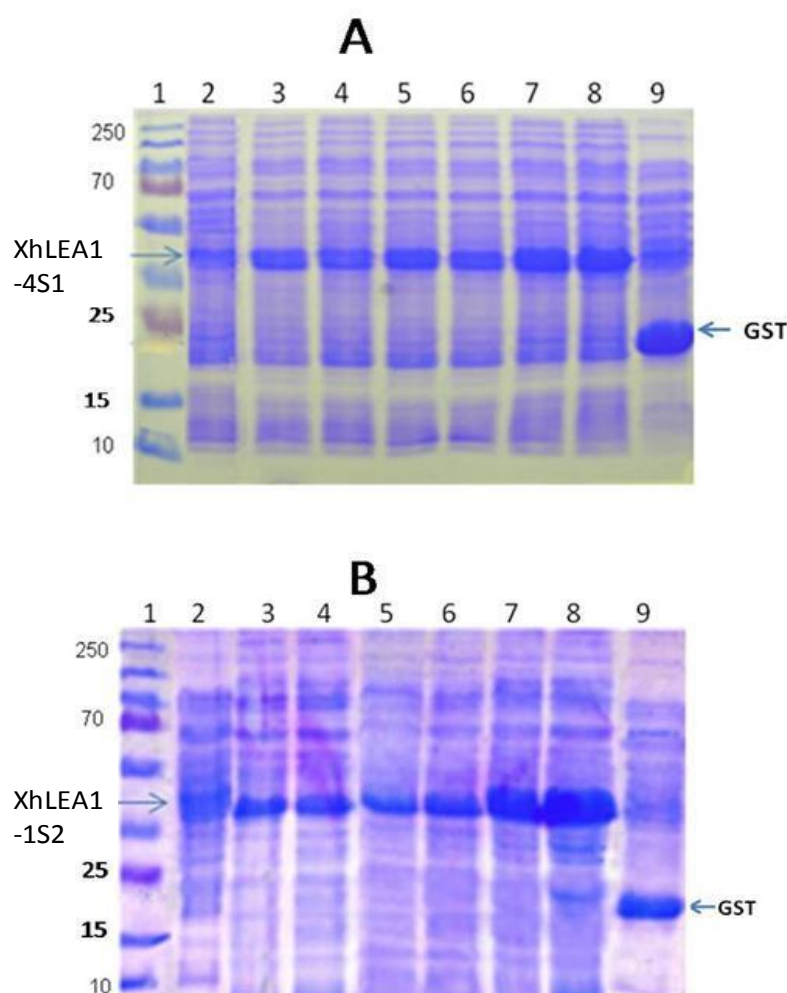


Figure 2.6 Optimization of *XhLEA1-4S1* (A) and *XhLEA1-1S2* (B) fusion protein expression conditions. Lane 1 in both A and B represent molecular weight marker (Prestained Protein Ladder, #SM1811); lane 2 is for un-induced samples, 3 – 8 represent 30, 60, 90, 120, 180 and 240 incubation times in minutes after IPTG induction. Strong band in lane 9 in both A and B represents GST protein produced from induced cultures containing empty pGEX-3X vector.

The GST-pGEX fusion proteins produced, as described above, had to be purified using a column of Glutathione Sepharose 4B beads that have affinity to the GST component of the fusion proteins. However, when elutes from the column were analysed, it was found out that the concentration of the proteins recovered was not satisfactory. The eluates contained a maximum of about 1.5 mg/ml protein when the glutathione column was reported to bind ≥ 8

mg/ml protein. Although purification of GST-LEA recombinant proteins using the GST purification kit resulted in a low yield of recombinant protein, the recombinant protein was pure, (Figure 2.7) and could be directly used in subsequent procedures without the need for further purification. Attempts were made to cleave the N-terminal GST from the XhLEA1 proteins with *Factor Xa*, but these were unsuccessful. Thus rabbits were injected with recombinant GST-XhLEA fusion proteins.

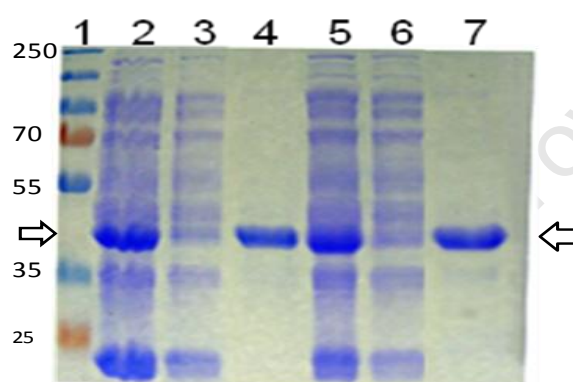


Figure 2.7 Purification of GST-*XhLEA1-4S1* and GST-*XhLEA1-1S2* fusion proteins. Lane 1 is molecular weight marker (Prestained Protein Ladder, #SM1811); lane 2 is total protein extract containing GST-*XhLEA1-4S1*, lane 3 is flow through and lane 4 is purified GST- *XhLEA1-4S1* protein. Lane 5 is total protein extract containing GST-*XhLEA1-1S2*, lane 6 is flow through and lane 7 is purified GST- *XhLEA1-1S2* . XhLEA-GST protein bands are indicated by arrows.

The identities of the purified *XhLEA1-4S1* and *XhLEA1-1S2* GST-fusion proteins injected into rabbits for the generation of antibodies were confirmed by means of mass spectrometry (MS) (Figure 2.8 and Figure 2.9). MS data analysis of both XhLEA1-4 and *XhLEA1-1S2* samples showed digestion patterns corresponding to the XhLEA proteins as well as the fusion protein. Tryptic digestion of GST-*XhLEA1-1S2* resulted in a spectrum with the most intense major peaks assigned to the fusion protein and the ions can be attributed to both LEA and GST fusion partners.

The ions assigned to *XhLEA1-4S1* were not the most intense. A number of the fragments also could not be assigned to this protein. This may be indicative of extraction artefacts or

modifications not accounted for (only deamidation modification was considered). Theoretical digestion of XhLEA1-4S1+GST would render 55 peptides including missed cleavages between $m/z = 800 - 3000$. Twenty three of the theoretical peptides were assigned experimentally observed ions with a correspondence of 42 %. The amino acid coverage of the detected peptides was 38 %. Theoretical digestion of the protein without the GST fusion resulted in 24 peptides of which 14 were detected experimentally for a correspondence of 54 %. The sequence coverage observed was 79 %.

Similarly *XhLEA1-1S2* +GST would theoretically render 46 peptides including missed cleavages between $m/z = 800 - 3000$. Twenty of the theoretical peptides were assigned experimentally observed ions with a correspondence of 43 %. The amino acid coverage of the detected peptides was 47 %. Theoretical digestion of this protein without the GST resulted in 14 peptides of which 10 were detected for a correspondence of 71 %. The sequence coverage observed was 57 %.

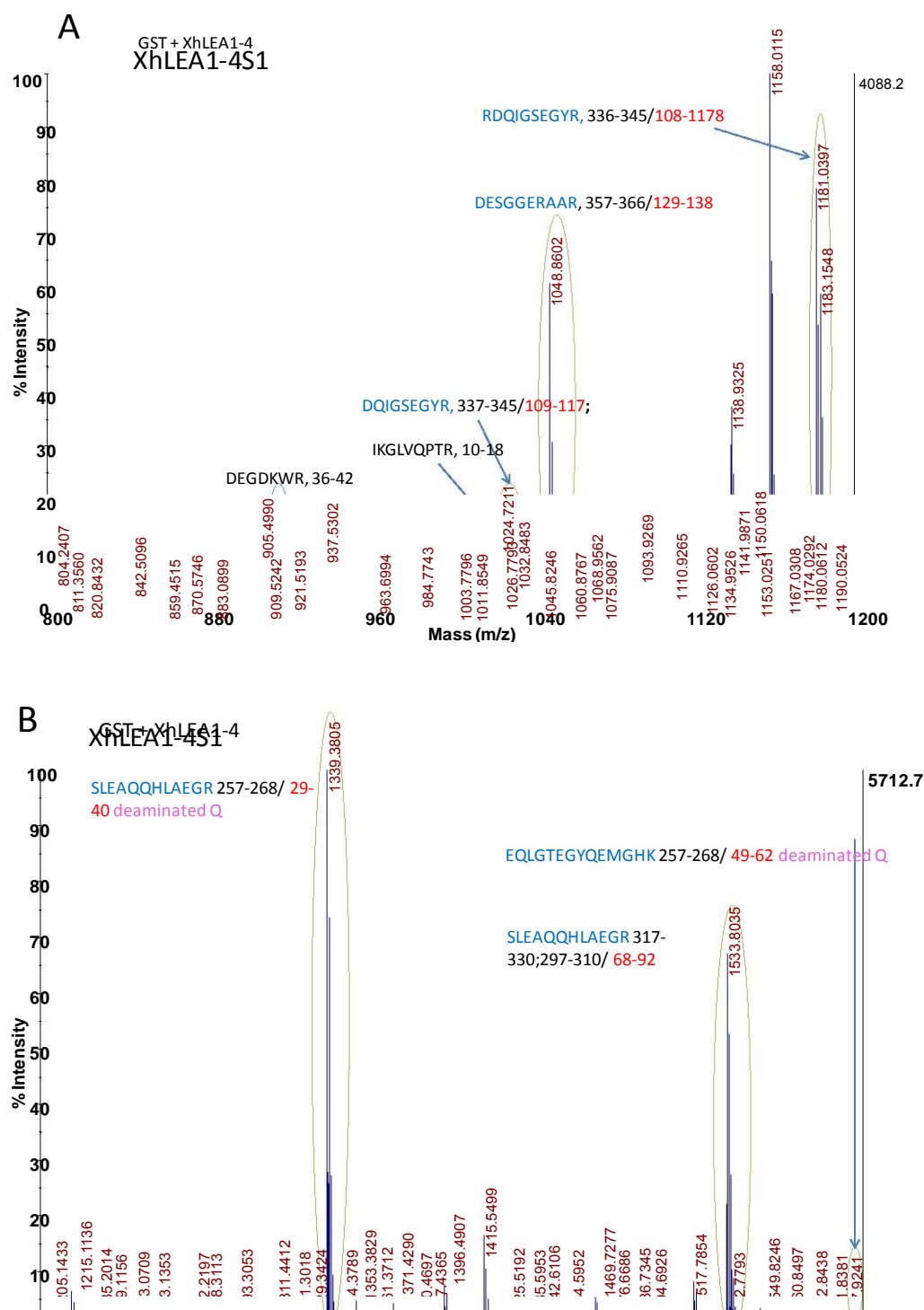


Figure 2.8 MS spectra of *GST+XhLEA1-4S1* (A and B) fusion protein, m/z= 800 - 1600. Ions corresponding to GST are indicated with blue circles and black text, and ions for *XhLEA1-4S1* protein with green circles and blue text. Deamidated ions are indicated.

Mass spectrum of XhLEA1-1S2 showing relative intensity versus mass-to-charge ratio (m/z). The spectrum displays several prominent peaks labeled with peptide sequences and their corresponding m/z values.

Key peaks identified:

- IELDKR (238-243/8-13) at m/z 801.4121
- DEGDKWR (36-42) at m/z 905.4976
- IKGLVQ²TR (10-18) at m/z 1011.8621
- IEAIPQIDK (192-190) at m/z 1039.2
- SVEAQENLAR (328-337/99-108) at m/z 1115.8343
- YEEHLYER (28-35) at m/z 1138.8301
- LILEYLEEK* (19-27) at m/z 1150.0513

The x-axis represents Mass (m/z) from 798.0 to 1200.0. The y-axis represents relative intensity from 0 to 100%.

2.3.4 Production and purification of Polyclonal antibodies

Indirect ELISA was utilized to determine the bleed with the highest titer and the optimal dilution of the polyclonal antibodies for antigen detection; antibodies from those bleeds were then further purified. The end point titer of the serum of both polyclonal antibodies for *XhLEA1-4S1* and *XhLEA1-1S2* was at a dilution of 10^6 . (A table showing the immune responses by rabbits to successive injections of GST-*XhLEA1-4S1* and GST-*XhLEA1-1S2* antigens is included in Appendix E).

Step wise purification of the serum involving saturated ammonium sulfate precipitation followed by filtration through DE-53 column removed non-IgG antibodies and other contaminating proteins from the serum. From the concentration measurements of the serial 1 ml elutes from the DE-52 column, it was determined that more than 90 % of the total serum protein was recovered in elutes 8 to 12 of a total of 15 ml (Table 2.3). However, the IgG antibody population recovered at this stage was made up of all the antibodies produced against the fusion protein and not specific to the LEA fusion partners only.

Table 2.3 Collection of IgG antibody serum fractions from DE52 column after precipitation with saturated ammonium sulphate.

Anti <i>XhLEA1-4S1</i>		Anti <i>XhLEA1-1S2</i>	
Elute number, 1 ml each	Concentration (mg/ml)	Elute number 1 ml each	Concentration(mg/ml)
1	0.03	1	0.02
2	0.04	2	0.07
3	0.08	3	0.02
4	0.03	4	0.01
5	0.04	5	0.02
6	0.01	6	0.02
7	6.05	7	0.07
8	11.64	8	9.35
9	12.89	9	14.70
10	15.25	10	16.06
11	11.41	11	16.06
12	1.38	12	14.98
13	0.57	13	5.55
14	0.38	14	0.35
15	0.19	15	0.26

NB: Concentration of eluates was measured using Nanodrop, $OD_{280} \times 14 = 1 \text{ mg/ml}$

The presence of heterogeneous IgG antibody population could be seen from the western blot results with total protein extracts from *E. coli* harbouring the pET21-a recombinant plasmids of *XhLEA1-4S1* or *XhLEA1-1S2* (Figure 2.10 A and B) (Note: the XhLEA proteins were also cloned into pET21-a vector as part of this PhD study for purposes described in Chapter 4). However, despite the non-specific signals, a good signal from the target proteins was detected. Furthermore, despite the 53 % amino acid similarity between *XhLEA1-4S1* and *XhLEA1-1S2* proteins that were as antigens, the cross reactivity observed was not significant.

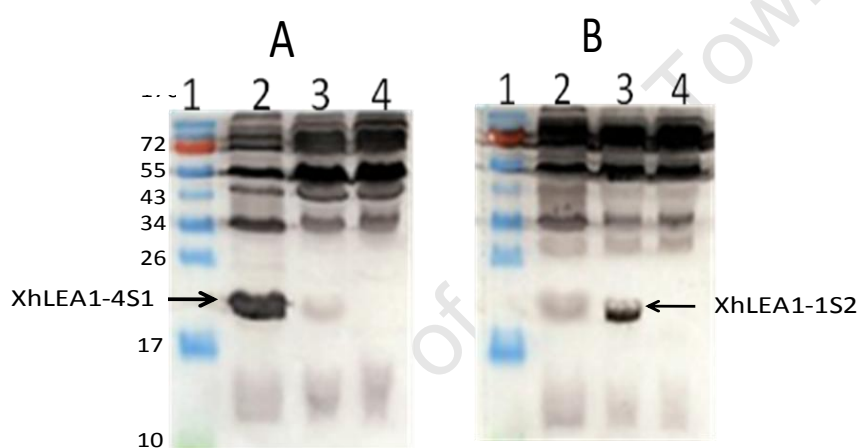


Figure 2.10 Checking the specificity and cross reactivity of anti *XhLEA1-4S1* (A) and *XhLEA1-1S2* (B) polyclonal antibodies (total IgG) using western blotting on recombinant proteins of these LEAs produced in pET-21(a+) vector. In both pictures lane 1 represent molecular weight marker (Prestained Protein Ladder, #SM0671); lane 2, total protein containing T7-tag-*XhLEA1-4S1*; lane 3, total protein containing T7-tag-*XhLEA1-1S2* protein, and lane 4 is total protein lysate extracted from *E. coli* containing empty vector (Refer Chapter 4 for T7-tagged XhLEA proteins).

The specificity of the polyclonal antibodies did not improve much with the use of GST-bound glutathione column targeting the removal of GST specific antibodies (results not shown). Although there was still some faint non-specific signals from both antibodies on films exposed for longer times, the acid elution purification method resulted in better quality of antibodies (Figure 2.11 A and B) with respect to the non-specific signals observed with total protein extracts (Figure 2.10 A and B) . The relative intensity of the signal from the

target proteins remained fairly the same with the use of both purified and un-purified antibodies. Although the titer of the antibodies after the purification steps was not reassessed, the optimal dilution of the purified antibodies (diluted in a 3 % fat free milk powder) that resulted in a good signal was found to be 1:500.

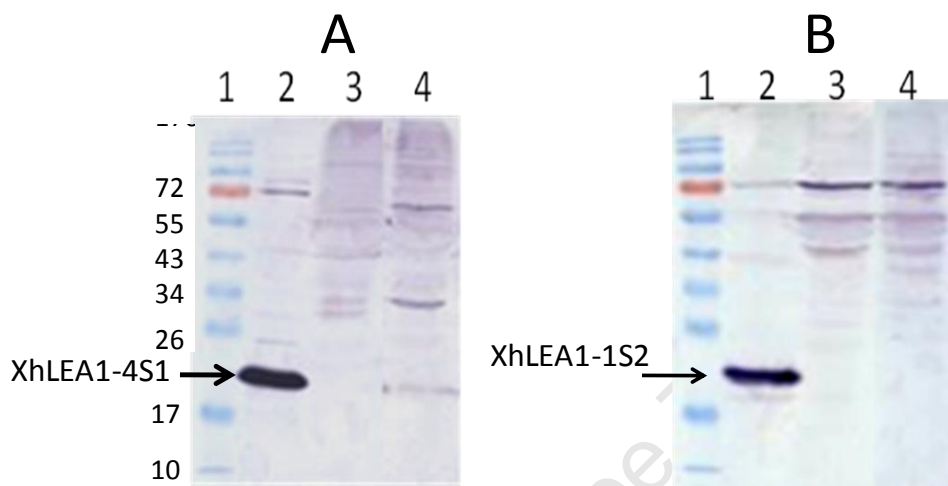


Figure 2.11 Checking the specificity of *XhLEA1-4S1* (A) and *XhLEA1-1S2* (B) antisera after acid elution purification. Lane 1 represent molecular weight marker in both A&B (Prestained Protein Ladder, #SM0671). Lane 2 represents *XhLEA1-4S1* (A) and *XhLEA1-1S2* (B) recombinant proteins expressed in pET21-a vector (Refer Chapter 4 for T7-tagged XhLEA proteins).. Lanes 3 and 4 are proteins extracted from fully hydrated and dry *X. humilis* leaves respectively.

2.3.5 *XhLEA1-4S1* and *XhLEA1-1S2* expression studies

Although different extraction methods were used to isolate total protein from various tissues of *X. humilis*, better results were obtained using the hydrophilic plant protein extraction kit. Protein extracted by this kit contained increased amounts of lower molecular weight proteins, i.e., in the range of the molecular weight of *XhLEA1-4S1* and *XhLEA1-1S2* compared to phenol or TCA (Trichloroacetic Acid) based extraction methods. Furthermore, specific and better expression signals were observed for the *XhLEA1-4S1* and *XhLEA1-1S2* proteins when the total protein extracted by the hydrophilic plant protein extraction kit was used, and hence the protein expression results presented below were obtained using this extraction kit.

Expression of both *XhLEA1-4S1* and *XhLEA1-1S2* was absent in hydrated leaves. During a desiccation cycle, expression of both LEA proteins was first detected at 55 % RWC, and remained constant until leaves were in the dry state (<10 % RWC, achieved after 12 days). Both these XhLEA proteins disappeared within 6 hours of rehydration of desiccated *X. humilis* leaves (Figure 2.12 A and B). There was also expression of the two XhLEA proteins in dry *X. humilis* seeds. Whereas, expression of *XhLEA1-1S2* in dry roots, dry leaves and seeds was detected when 20 µg protein was loaded, no signal was observed for *XhLEA1-4S1* in dry roots (Figure 2.12 B). A strong signal was also detected in the protein extract of *Arabidopsis* seeds using the anti-*XhLEA1-4S1* antibody. The corresponding protein for the latter signal was estimated to be around 17 kDa (Figure 2.12 A). The same results were obtained in a further 2 independent experiments.

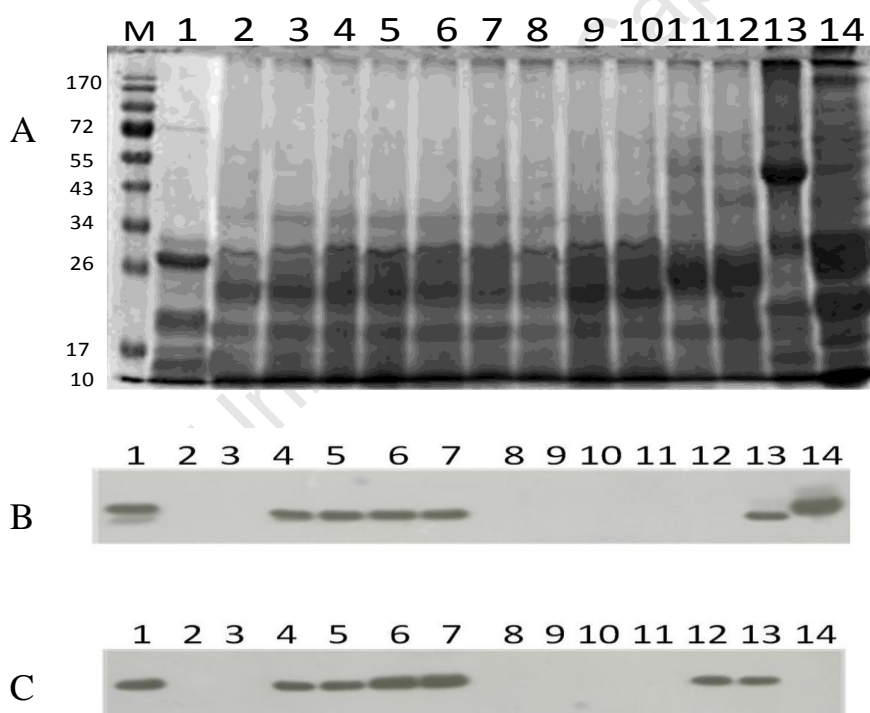


Figure 2.12 Western blot expression studies of *XhLEA1-4S1* (panel B) and *XhLEA1-1S2* (panel C) proteins using polyclonal antibodies. Panel ‘A’ shows a replicate (control) SDS gel picture of the proteins used for the western. 20 µg protein from each of the following preparations was loaded in each lane. Lane1 = positive controls for *XhLEA1-4S1* in (A) and *XhLEA1-1S2* in (B); Lane 2= fully hydrated leaves, Lanes 3 – 6 for leaf samples at 75, 55, 35 and 20 % RWCs respectively, and Lane 7 for fully dry leaves. Lanes 8, 9 and 10 are for 6, 18 and 42 hours re-hydrated plant leaves; Lanes 11 and 12 are for fully hydrated and dry roots respectively. Lanes 13 and 14 are for *X. humilis* and *Arabidopsis* seed proteins respectively. Experiment was repeated 3 times and the same results were obtained.

From previous gene expression study where up-regulation of mRNA of both *XhLEAs* in roots and leaves of *X. humilis*, in response to desiccation, was reported (Ngubane 2008), it was suspected that the *XhLEA1-4S1* protein was also present in dry roots of this plant. Therefore, increased amounts of protein extracts from dry roots were used in another round of western blot study and protein transfer was checked by Ponceau staining (Figure 2.13 A). An expression signal for *XhLEA1-4S1* was detected when the amount of protein loaded in the initial study was doubled to 40 μ g, and stronger signal was observed from 50 μ g of dry root protein extract (Figure 2.13 B).

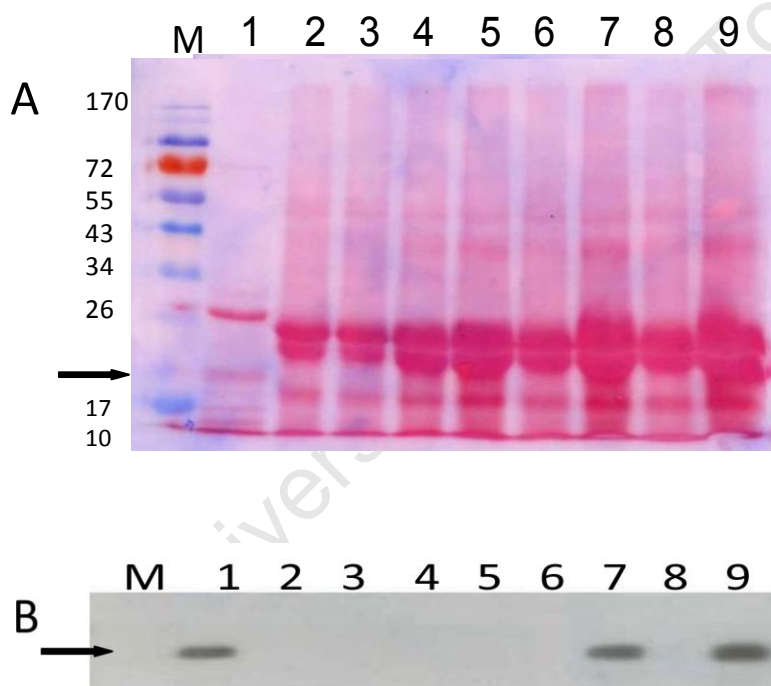


Figure 2.13 Western blot expression study of *XhLEA1-4S1* in dry roots. In ‘A’ (Ponceau staining), lane M is molecular weight marker (#SM0671); lane 2 is positive control (partially purified recombinant protein of *XhLEA1-4S1*); lanes 2, 4, 6, and 8 were loaded with 20 μ g protein extracted from hydrated *X. humilis* roots (negative control); lanes 3, 5, 7 and 9 were loaded with 20, 30, 40 and 50 μ g of dry root protein, respectively. Western blot results are shown in ‘B’. Expression of *XhLEA1-4S1* was detected in 40 and 50 μ g of root protein extract.

2.4 Discussion

2.4.1 Sub-cloning of *XhLEA1-4S1* and *XhLEA1-1S2* cDNAs and production of recombinant proteins

As measured first by the growth of colonies on selective media and then by colony PCR of the recombinant clones, the sub-cloning of *XhLEA1-4S1* and *XhLEA1-1S2* into pGEX-3X expression plasmid was successful. Growth of even a single colony with the correct recombinant plasmid was sufficient for the intended purpose. Furthermore, by comparing the growth of the cells expressing the recombinant GST-XhLEA proteins with those that did not have the recombinant plasmids, it was demonstrated that expression of the recombinant fusion proteins did not have deleterious effects on the growth of *E. coli*.

Elution results *XhLEA1-4S1* and *XhLEA1-1S2* fusion proteins using glutathione beads were more dilute than would be expected of glutathione column saturated with GST protein. Such a column is reported to have a capacity to bind ≥ 8 mg/ml GST protein. It has been pointed out that the GST fusion proteins may not behave as non-fused GST in the induction and purification stages and that optimization would be required for experiment involving the GST system (Mercado-Pimentel *et al.*, 2002). It was likely that the glutathione 4B resin was not saturated with the fusion protein resulting in lower protein yield when final elutes were analysed.

Poor results of GST fusion proteins due to lower protein solubility were reported before (Mercado-Pimentel *et al.*, 2002). Slow binding between glutathione and GST beads was also reported to cause low protein yield (Glutathione Resin User Manual, 2009; Nieslanik and Atkins, 2000). Protein solubility could improve by adding solubilisers such as Sarkosyl and Triton X-100 (Frangioni and Neel, 1993; Tao *et al.*, 2010). In the current study, longer incubation of the recombinant proteins with the glutathione 4B resin and slow flow rate

improved yield considerably. Longer incubation time of the GST protein with the beads and lower flow rate might have helped the beads to bind to more quantities of fusion protein.

Confirming the identity of the GST-XhLEA fusion protein was required before these proteins were administered to rabbits to generate antibodies. This was done to make sure that the protein of interest was used for the production of the required antibodies. The band corresponding to differentially expressed proteins could be located if separated on SDS-PAGE along with control and a molecular weight marker. Knowledge of the predicted amino acid sequence was also important in locating the protein band of interest on SDS-PAGE. However, MS spectrometry was a better tool in confirming the identity of the GST-XhLEA proteins as it showed the amino acid composition of the proteins. The difference observed between experimental and theoretical MS results was not unexpected as the ideal experimental conditions cannot be achieved in practical setting. However, the MS results obtained provided sufficient data to confirm the identity of the GST-XhLEA fusion proteins.

2.4.2 Production and purification of polyclonal antibodies

GST-fused XhLEA proteins were used for the generation of polyclonal antibodies as efforts to cleave off the N-terminal GST did not succeed. It is possible that the big GST protein folded in a way that masked the cleavage site preventing the Factor *Xa* enzyme access to act. Failure to remove GST was not a big problem for the current study as the XhLEA-specific antibodies were able to be isolated from the anti-GST-XhLEA antibody population. However, for other functional studies, other expression vectors that do not add bigger fusion tags to the recombinant proteins of interest are recommended (e.g. T7 or His-Tags).

The size of *XhLEA1-4S1* or *XhLEA1-1S2* proteins without the GST partner was near the size of protein antigens (≤ 10 kDa) that require conjugation or cross-linking of immunogenic

protein or carrier molecule that increases immunogenicity of small antigens. This is due to the fact that smaller sized antigens do not cause vigorous immune response required for the production of antibodies when injected into host organisms, instead they induce tolerance (Hanly *et al.*, 1995). Due to the scope of the current study, the antigenicity of either the XhLEA proteins or the GST protein was not investigated separately, but since the polyclonal antibodies produced against the fusion protein reacted to both fusion partner antigens, the GST-XhLEA fusion protein must have good antigenic properties for the production of antibodies in rabbits.

The other aspect of protein antigens that has to be considered is their secondary structure. Globular and aggregated proteins are reported to be better antigens (Hanly *et al.*, 1995; Hancock and O'Reilly, 2004). Structural investigation of the amino acid sequences of *XhLEA1-4S1* and *XhLEA1-1S2* indicated that these LEA proteins, like many other LEAs, are unstructured and hence do not have the qualities required for good immunogenicity. Therefore, the vigorous immune response by the rabbits to the antigens used in the current study could be an indication that the GST fusion partner had positive effect on the antigenicity of the XhLEA recombinant proteins.

Ammonium sulphate precipitation step used in the antibody purification method facilitated further immunoaffinity based antibody purification steps as it effectively separated IgG antibodies from other serum proteins. The use of this step in the purification of monoclonal antibodies, where highly purified and known antigen is used, results in relatively pure antibodies that could be used directly without the need of further purifications. The subsequent affinity based purification for the removal of GST specific antibodies would work better if the amount of GST specific antibodies in the sample as well as the binding kinetics between GST and the glutathione beads was known, i.e. knowing the amount of antibodies would help in binding sufficient amounts of GST to the beads.

The non-specific signals observed initially in the western blot experiments might not have resulted from a strong affinity of the polyclonal antibodies to the non-target proteins. If the latter was the case then the signal would persist despite the continued purification procedures performed. The observed non-specific signals could have been due to the presence of other contaminating antibodies produced as a result of the use of non-pure antigen at the rabbit immunization step. Although, MS results showed that the proteins (antigens) used were GST-*XhLEA1-4S1* or GST-*XhLEA1-1S2*, these proteins were not extracted from a single band as for the MS analysis. Instead the corresponding recombinant proteins were purified using GST affinity column. These samples might have contained smaller amounts of contaminating proteins that somehow escaped the purification steps resulting in the production contaminated antibodies. Although the purification procedure was labour intensive, I was able to generate pure polyclonal antibodies against *XhLEA1-4S1* and *XhLEA1-1S2* recombinant proteins.

2.4.3 *In planta* expression studies

The failure to detect expression of XhLEA proteins from crude extracts, extracted by phenol or TCA based methods, could be related to the concentration of the respective XhLEAs compared to other constitutively expressed hydrophilic proteins; and if these proteins are indeed expressed in such low quantities, it might imply that their role in the tissues studied does not require expression of large amounts of proteins. Alternatively, the biological function of such proteins might require expression of other similar proteins or osmotic adjustment strategies such as increased levels of compatible solutes as has been previously suggested (Maitra and Cushman 1994).

The hydrophilic plant protein extraction kit must have enriched the concentration of the XhLEA proteins along all the other hydrophilic proteins. This extraction method is helpful when working with proteins whose hydrophathic properties are known. The drawback of the approach is that it might overestimate the expression of target proteins when they are actually produced in very low quantities under physiological conditions. However, for studies like the current one, conclusions can be drawn from comparative analysis of results obtained from control and experimental samples extracted using the same procedure.

The expression of *XhLEA1-4S1* and *XhLEA1-1S2* proteins in leaves of *X. humilis* was induced at and below 55 % RWC. It is thus possible that the 60 to 50 % RWC range during dehydration is critical for desiccation tolerant plants to put protection mechanisms in place facilitated by these LEAs. This suggestion was supported by earlier findings that the viability of soybean seedlings was not affected until the RWC dropped below 60 % (Whitsitt *et al.*, 1997). Expression results also showed that *XhLEA1-4S1* protein was less abundant in dry roots compared to dry leaves and seeds. This type of tissue specific variation of Group 1 LEA protein expression was reported previously where decreased expression was observed in embryonic axis of horsegram (*Macrotyloma uniflorum* (Lam.) Verde.)Cv.VZM1) compared to cotyledons in 24 hr stressed seedlings (Veeranagamallaiah *et al.*, 2010).

XhLEA1-4S1 and *XhLEA1-1S2* proteins were not detected in rehydrating tissues demonstrating that their expression is not required under conditions where shortage of water is not a problem. These results generally agree with previous report where the mRNA expression levels of these XhLEA genes dropped on rehydration (Ngubane, 2008). Similar down regulation of LEA mRNA levels was also reported on rehydrating seeds of *Ricinus communis* L. (Han *et al.*, 1997).

XhLEA1-4S1 and *XhLEA1-1S2* proteins were also expressed in seeds presumably for the same purpose; that is to prevent damage from desiccation. Possible explanation on how these proteins might play a protective role against desiccation stress is presented in Chapter 4. The *XhLEA1-4S1* antibody also recognized a protein present in dry seeds of *A. thaliana* which was estimated, from its migration on SDS-PAGE, to be around 17 kDa in size. Based on the amino acid sequence similarity shared between *XhLEA1-4S1* and the *EM1* protein from *Arabidopsis* (*AtEm1*, predicted molecular weight: 16.6 kDa; Bies *et al.*, 1998), the protein detected could be *EM1*.

One of the objectives of this PhD study, the subject of this chapter, was to characterize the expression conditions of *XhLEA1-4S1* and *XhLEA1-1S2* in *X. humilis*. The results obtained showed that these XhLEA proteins, whether it was in dry leaves or roots or in seeds of *X. humilis*, were expressed in response to desiccation stress and are likely to have a protective role against damage from desiccation. These results pave the way forward to perform *in vitro* functional studies on *XhLEA1-4S1* and *XhLEA1-1S2* proteins; and transformation experiments where the respective genes could be expressed in other organisms such as yeast or *Arabidopsis* for *in vitro* function testing.

Chapter 3

Immunolocalization of *XhLEA1-4S1* and *XhLEA1-1S2* proteins in *X. humilis* leaves, roots and seeds

3.1 Introduction

Knowledge of intracellular location of a protein is important in determining its function. For instance, if a desiccation responsive LEA protein is found associated with membranes, then its function is likely to be membrane stabilization; and if found associated with nuclei then its function is likely to be DNA stabilization. Localization studies of LEA proteins in plants and in anhydrobiotic organisms indicate that they occur in a variety of tissues and cellular organelles under different stress conditions; however some are specific to certain stresses and cell or tissue types (Nylander *et al.*, 2001; Karlson *et al.*, 2003; Rorat *et al.*, 2006; Shih *et al.*, 2010b).

LEA gene expression studies in seeds using *in situ* hybridization technique, at the tissue level, revealed that the transcripts of most of the hydrophilic *LEA* genes investigated accumulate in vascular tissues of the cotyledon and embryo axis (epicotyl and radical; Shih *et al.* (2010b)). One example, a soybean (*Glycine max* L.) Group 1 LEA transcript, *GmPM11* (accession no. AF004805) that encodes an *Em6*-type protein, was found to accumulate in vascular tissues of radicle of fresh soybean seeds (Shih *et al.*, 2010b).

As indicated in the table below, subcellular localization studies indicate the occurrence of various LEAs in structures including membranes, mitochondria, vacuoles, endoplasmic reticulum, cytoplasm and nucleus (Borovskii *et al.*, 2002; Heyen *et al.*, 2002; Shih *et al.*, 2010b; Brini *et al.*, 2007a; Marttila *et al.*, 1996; Zhao *et al.*, 2010; Colmenero-Flores *et al.*, 1999; Kalifa *et al.*, 2004a; Rorat *et al.*, 2006). It was also found out that different groups of LEAs occur in more than one cellular compartment (Battaglia *et al.*, 2008; Hand *et al.*,

2010; Shih *et al.*, 2010b; Tunnacliffe and wise, 2007). The presence of LEAs associated with many cellular structures or tissue types suggests that their presence is required by all these structures during water stress. Thus far, the subcellular location of about 20 LEA proteins mainly from Group 2 and 3 have been determined experimentally (Table 3.1, taken from Tunnacliffe and Wise, 2007).

Table 3.1 LEA proteins with known subcellular localisation (From Tunnacliffe and Wise 2007).

ID	LEA Group	Species	Location	Reference
DHN1_MAIZE	2	<i>Zea mays</i> (maize)	Nuc/Cyt	Godoy <i>et al.</i> 1994
ERD14_ARATH	2	<i>Arabidopsis thaliana</i> (mouse-ear cress)	Clp/Per	SUBA At1g76180
DHR18_ARATH	2	<i>Arabidopsis thaliana</i> (mouse-ear cress)	Nuc/Cyt	Nylander <i>et al.</i> 2001
COR47_ARATH	2	<i>Arabidopsis thaliana</i> (mouse-ear cress)	Nuc/Cyt	SUBA At1g20440
Q9ZR21_CITUN	2	<i>Citrus unshiu</i> (satsuma orange)	Mit ^a	Hara <i>et al.</i> 2003
TAS14_SOLLC	2	<i>Solanum lycopersicum</i> (tomato)	Nuc/Cyt	Godoy <i>et al.</i> 1994
DHR21_ORYSA	2	<i>Oryza sativa</i> (rice)	Cyt	Mundy and Chua 1988
CS120_WHEAT	2	<i>Triticum aestivum</i> (wheat)	Nuc/Cyt	Houde <i>et al.</i> 1995
CO410_WHEAT	2	<i>Triticum aestivum</i> (wheat)	PlMem	Danyluk <i>et al.</i> 1998
VCaB45 ^b	2	<i>Apium graveolens</i> (celery)	Vac	Heyen <i>et al.</i> 2002
LEA13_GOSHI, LEAD7_GOSHI	2,3	<i>Gossypium hirsutum</i> (cotton)	Cyt	Roberts <i>et al.</i> 1993
Q06540_WHEAT	3	<i>Triticum aestivum</i> (wheat)	Clp	NDong <i>et al.</i> 2002
Q8S385_SECCE	3	<i>Secale cereale</i> (rye)	Clp	NDong <i>et al.</i> 2002
DRPF_CRAPL	3	<i>Craterostigma plantagineum</i> (res. plant)	Clp	Iturriaga <i>et al.</i> 1992
Q42512_ARATH	3	<i>Arabidopsis thaliana</i> (mouse-ear cress)	Clp	Lin and Thomashow 1992
LEAD8_DAUCA	3	<i>Daucus carota</i> (carrot)	Cyt	Franz <i>et al.</i> 1989
Q39873_SOYBN	3	<i>Glycine max</i> (soybean)	ER	Hsing <i>et al.</i> 1995
LEA1_HORVU	3	<i>Hordeum vulgare</i> (barley)	PSV/Cyt	Marttila <i>et al.</i> 1996
Q93Y63_9ROSA	3	<i>Morus bombycis</i> (mulberry tree)	ER	Ukaji <i>et al.</i> 2001
Q41060_PEA	3	<i>Pisum sativum</i> (garden pea)	Cyt	Alban <i>et al.</i> 2000
Q5NJL5_PEA	3	<i>Pisum sativum</i> (garden pea)	Mit	Grelet <i>et al.</i> 2005
LEA1_APHAV	3	<i>Aphelenchus avenae</i> (nematode)	Cyt	Goyal <i>et al.</i> 2005

Clp: Chloroplast, Per: peroxisome, Nuc: nucleus, Cyt: cytoplasm, Mit: mitochondria, PlMem: plasma membrane, ER: endoplasmic reticulum/secreted, PSV: protein storage vacuole, Vac: vacuole. SUBA (for the Subcellular Location of Proteins in *Arabidopsis* Database) is followed by the TAIR gene name corresponding to the respective *Arabidopsis thaliana* LEA proteins (Heazlewood *et al.* 2005).

^a In transgenic tobacco

^b The sequence is given in Heyen *et al.* (2002).

Currently, the two main methods used for protein localization studies are microscopy and bioinformatics computer predictor programs. Whereas the predictors examine protein sequences for the presence of N-terminal signals (Small *et al.*, 2004) that associate these proteins with specific cellular structures, microscopy employs techniques such as immunogold labelling of target proteins (antigens) in fixed tissues and use of protein-

specific antibodies to locate them; or tagging proteins with fluorescent material and then following their localization by confocal microscopy (Heazlewood *et al.*, 2005).

Immunogold-labelling technique was used in the current study to determine the subcellular localization of *XhLEA1-4S1* and *XhLEA1-1S2* proteins in *X. humilis* leaves, roots and seeds. A similar localization study was reported previously where antiserum raised against a *Craterostigma plantagineum* protein was used to localize three Group 2 LEA proteins in various subcellular compartments in *Betula pubescens* (birch) (Rinne *et al.*, 1999).

Immunogold-labelling is a useful method for the visualization of the sites of biological molecules or structures within cells or tissues (Roth *et al.*, 1978). This technique involves Transmission Electron Microscopy (TEM) that is suitable for localization and counting of gold particles associated with target antigens at an ultrastructural level. Tissue samples are resin-embedded or cryofixed and sectioned. Gold-conjugated antibodies are then used to detect and localize target antigens on ultrathin sections under the microscope (Mayhew, 2005).

Quantitative and qualitative techniques are used to determine the location of the antigen as well as its spatial distribution pattern in different intracellular compartments of the same sample of cells or between experimental groups (Griffiths and Hoppeler, 1986; Mayhew *et al.*, 2004). The method of analyses of the distribution of gold labels depends on the method of tissue fixation used, the nature of the cell compartments studied and the purpose of the study. Thus, different approaches have been described for the quantitative evaluation of gold label distribution patterns (Griffiths *et al.*, 2001; Mayhew *et al.*, 2004; Mayhew 2005).

Mayhew (2005) described a method for the quantitative evaluation of immunogold labelling convenient for use with different types of cellular compartments. This method does not

require standardization of magnification within and across groups provided that gold labels are visible for counting. In using this method, gold particles in the compartments are counted and their distribution compared by contingency table analysis with degrees of freedom determined by the number of compartments and experimental groups. Presence or absence of preferential gold labelling is then determined by the total χ^2 -value. However, the sizes and shapes of cellular compartments might change as a result of treatments and hence become non-comparable. An example where such structural change is observed is in the vacuoles of resurrection plants where the large vacuole in hydrated tissues is divided into smaller vacuoles in dry tissues. In such cases, additional considerations have to be made in order to compare the results observed.

Comparison of labelling at the organelle level using contingency tables also does not tell the label distribution pattern within a single compartment and hence does not distinguish between specific and non-specific labelling. Therefore, one should use additional analysis techniques to investigate this. Point pattern analysis techniques such as the nearest neighbour analysis (D'Amico and Skarmoutsou, 2008) can be used for this purpose (described in section 3.2.4). Following the confirmation of the up-regulated expression of *XhLEA1-4S1* and *XHLEA1-1S2* in dry leaves and roots, and also in seeds (Chapter 2), immunogold labelling was used to investigate the subcellular localization of these proteins in order to shed light on their function on desiccation tolerance.

3.2 Materials and Methods

3.2.1 Tissue fixation

Growth and treatment of *X. humilis* plants were as described in Chapter 2. Tissues from fully hydrated and dry leaves and roots as well as seeds were fixed as described by Sherwin and Farrant (1998). Several 1 mm square sections were cut from the mid-section of 3 randomly

selected fully hydrated and dry leaves and main roots each and fixed for 16h at 4 °C in 4 % paraformaldehyde in 0.1M phosphate buffer (pH 7.4) with 0.5% caffeine; and dehydrated with an increasing ethanol concentration (30 to 100 %). The samples were then gradually infiltrated with LR white resin and embedded in 100 % LR white resin (Spurr, 1969). The seed coat was removed from the seeds to allow the resin to penetrate, and then the seeds were fixed, dried and embedded as above.

3.2.2 Immunogold labelling and electron microscopy

95 nm thick sections of embedded samples were cut using Reichert Ultracut-S microtome (Reichert, <http://www.reichert.com>) and mounted on nickel grids coated in formvar. Samples were blocked in 1.5 % Tween, 1 % BSA and 1 % glycine in PBS for 1 hr, washed 3 times with PBS, 3 minutes each, and then transferred to either Anti-*XhLEA1-4S1* or Anti-*XhLEA1-1S2* polyclonal antibody dilution in PBS containing 1 % BSA and incubated for 3 hours at room temperature. Control grids were placed on 1 % BSA in PBS with no primary antibody added. After 3 hours, samples were washed 7 times with PBS containing 0.05 % Tween for 10 minutes each. Samples were incubated for 30 minutes in 1:50 dilution of 15-nm gold-conjugated goat-anti rabbit secondary antibody (Sigma) in 1 % BSA solution. Samples were washed 6 times with PBS containing 0.05 % Tween and 3 times with PBS only and then fixed in 2 % glutaraldehyde in PBS for 4 minutes, after which they were washed 5 times with double distilled water. The grids were stained with 2% uranyl acetate and 1% lead citrate (Reynolds, 1963) for 10 minutes each. Samples were examined with a LEO912 transmission electron microscope and images were recorded.

3.2.3 Counting of immunogold labels

TEM micrographs of 3 randomly selected cells each from hydrated and dry samples, 30 pictures per cell, were patched together for structure identification. Cell structures were then identified and label within them counted. In *X. humilis* and in many of the resurrection

plants studied, the large vacuole found in hydrated cells is split into smaller vacuoles in the desiccated state (Farrant 2000; Hedderson *et al.* 2009). Therefore, the mean area of the dry vacuoles was calculated and proportional area of the vacuoles in the hydrated state was considered for label counting. Attempts were made to use Image-Pro and ImageJ softwares to count gold labels in the images. However, due to the grainy nature of the TEM pictures, these programs were found to be not suitable. Gold labels in chloroplasts (n=15), vacuoles (n=15), mitochondria (n=15) were thus separately counted manually. As it was difficult to distinguish plasma membrane from cell wall, the mean number of gold labels found in cell wall-plasma membrane area of three cells was used. Similarly, labels within the cytosol and all un-identified structures were considered together and termed ‘cytosol and rest of the cell.’

3.2.4 Analysis of immunogold labelling data

Counts of observed gold labels were used to generate expected values for each corresponding compartment as described by Mayhew (2005) and Mayhew *et al.* (2004). For each compartment, the number of expected particles is calculated by multiplying the corresponding column sum by the corresponding row sum divided by the grand row sum. The partial χ^2 value for each row was also determined as follows: $\chi^2 = (\text{observed} - \text{expected})^2 / \text{expected}$; and the sum of all the partial χ^2 values was added to give a total χ^2 value. This value with the corresponding degrees of freedom was used to determine whether there is a preferential labelling or not. Immunogold labelling in compartments that showed preferential labelling was further investigated by nearest neighbour analysis using the pair correlation function (PCF), $g(r)$ (Gathercole *et al.*, 2000; D’Amico and Skarmoustsou 2008).

$$g(r) = n(r) / 2\rho\pi r\Delta r,$$

where $n(r)$ is the mean number of gold particles in a circle with width Δr at a distance r and ρ is the gold particle density. $g(r) > 1$ indicates clustering, $g(r) < 1$ indicates regular distribution, and $g(r)=1$ random distribution. Population $g(r)$ was determined by randomly

selecting areas of $642.86\mu\text{m}^2$ and constructing concentric circles of 50 nm width around 10 different reference gold particles from electron micrographs of samples (Gathercole *et al.*, 2000; D'Amico and Skarmoustsou 2008). Label observed to form distinct clusters (of > 5 spots) all within a given cell compartment were distinguished from individual, more randomly located spots. The former were considered to indicate significant sites of labelling and the latter the result of non-specific labelling.

3.3 Results

Identification of cellular structures from individual TEM pictures, 30 micrographs of each cell, taken at 1000 nm scale was done by aligning these pictures and reconstructing the whole cell. Although some detachment of the cytoplasm from the walls was observed in tissues of both hydrated and dry samples, the fixation method was good enough for the identification of cellular structures and gold particle counting. Furthermore, cell walls in dry tissues were seen folded as reported in other resurrection plants (Vicre *et al.*, 1999) and many of these were darkly stained. However, gold labels were clearly identifiable for counting. Note that the white spots on pictures were formed as a result of the holes created during coating of the nickel grids by fomvar.

Only one or two gold labels were observed in the negative controls where no primary antibody was added, and hence these 'no-primary-antibody' samples were not considered for any of the analysis performed. Instead, hydrated leaves and roots [(which showed no expression either by RT-PCR analysis (Ngubane, 2008) or by Western blot analysis (Chapter 2)], which were treated with primary and secondary antibodies, were used as negative controls. Labelling was observed in varying amounts in the sections of leaves and roots of both fully hydrated and dry samples (Figures 3.1-3.8) as well as seeds (Figures 3.9 & 3.10.). Whereas labelling in the plasma membrane and cell wall areas of dry leaf and dry root cells showed a clustering pattern of gold labels (arrowed), labelling in the rest of the

cells for both hydrated and dry samples was random. Representative electron micrograph pictures from leaf and root tissues are shown below.

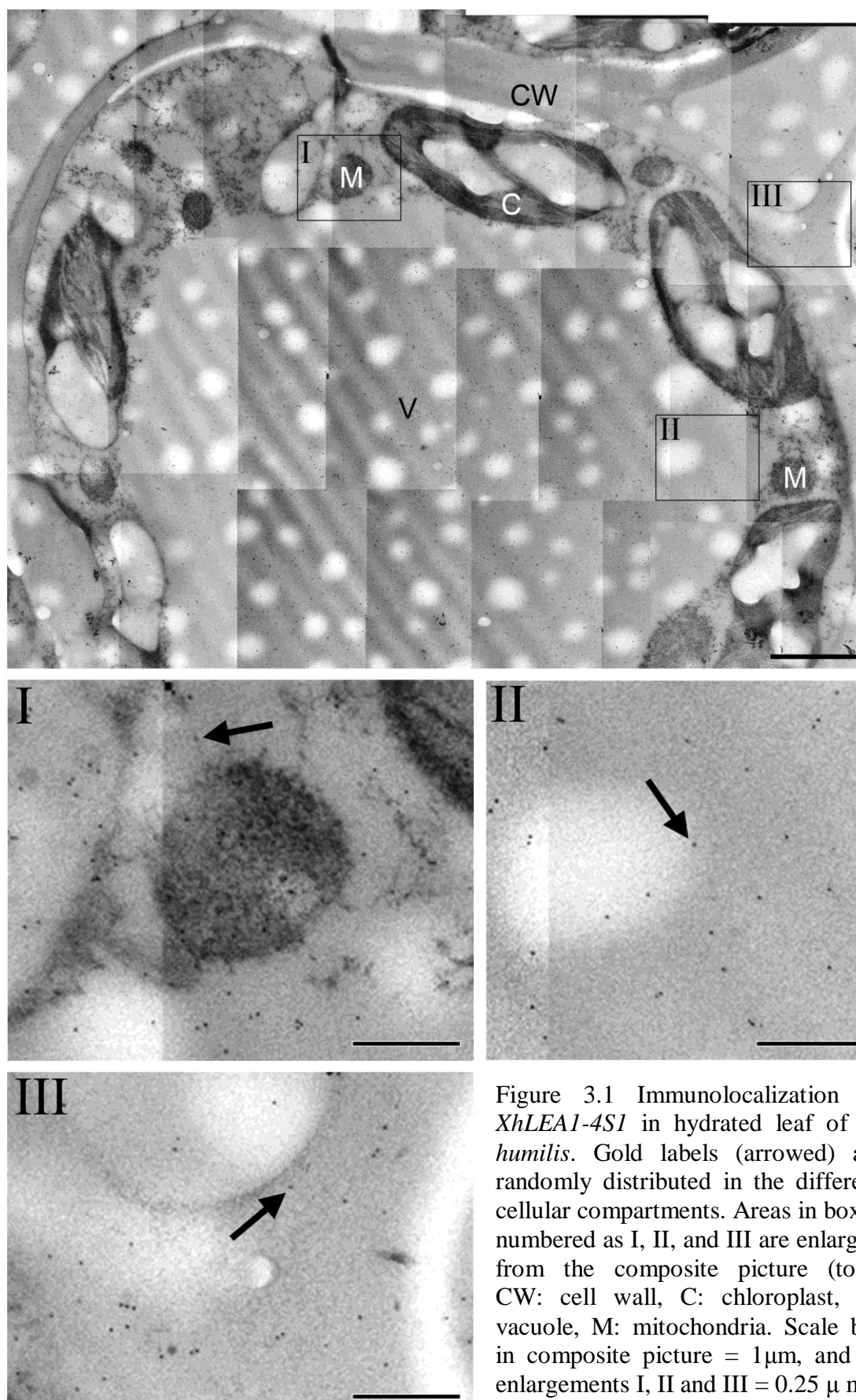
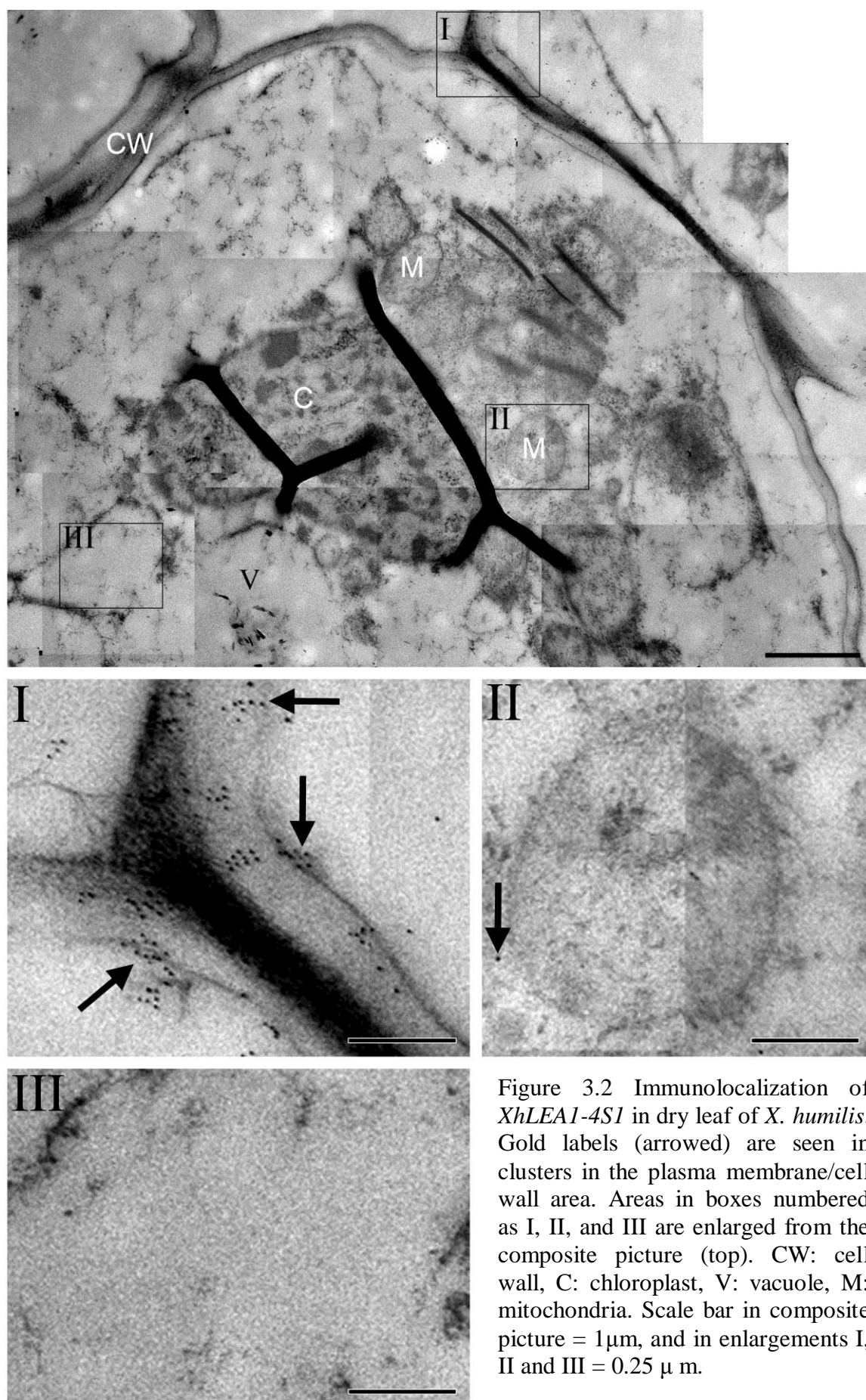
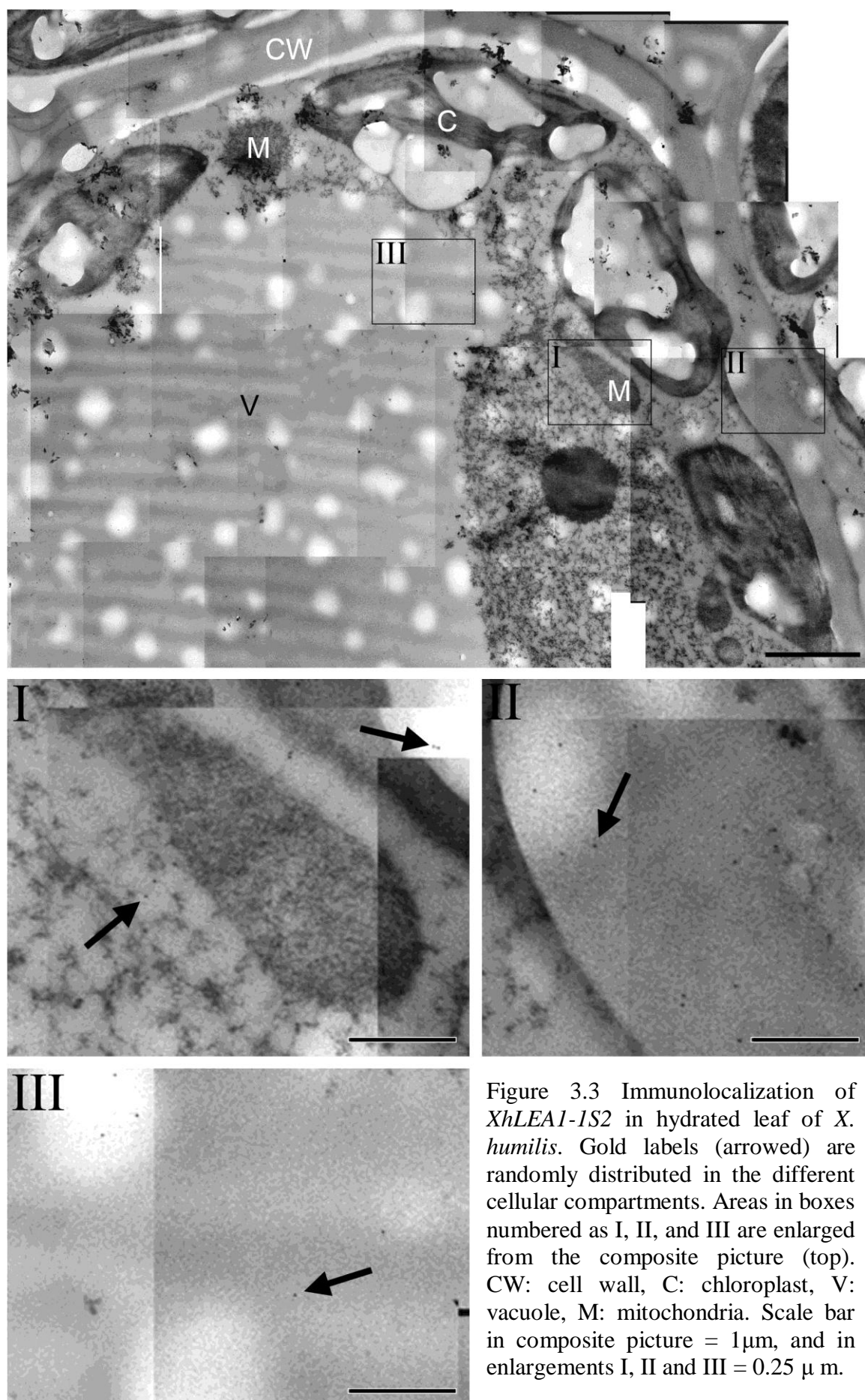
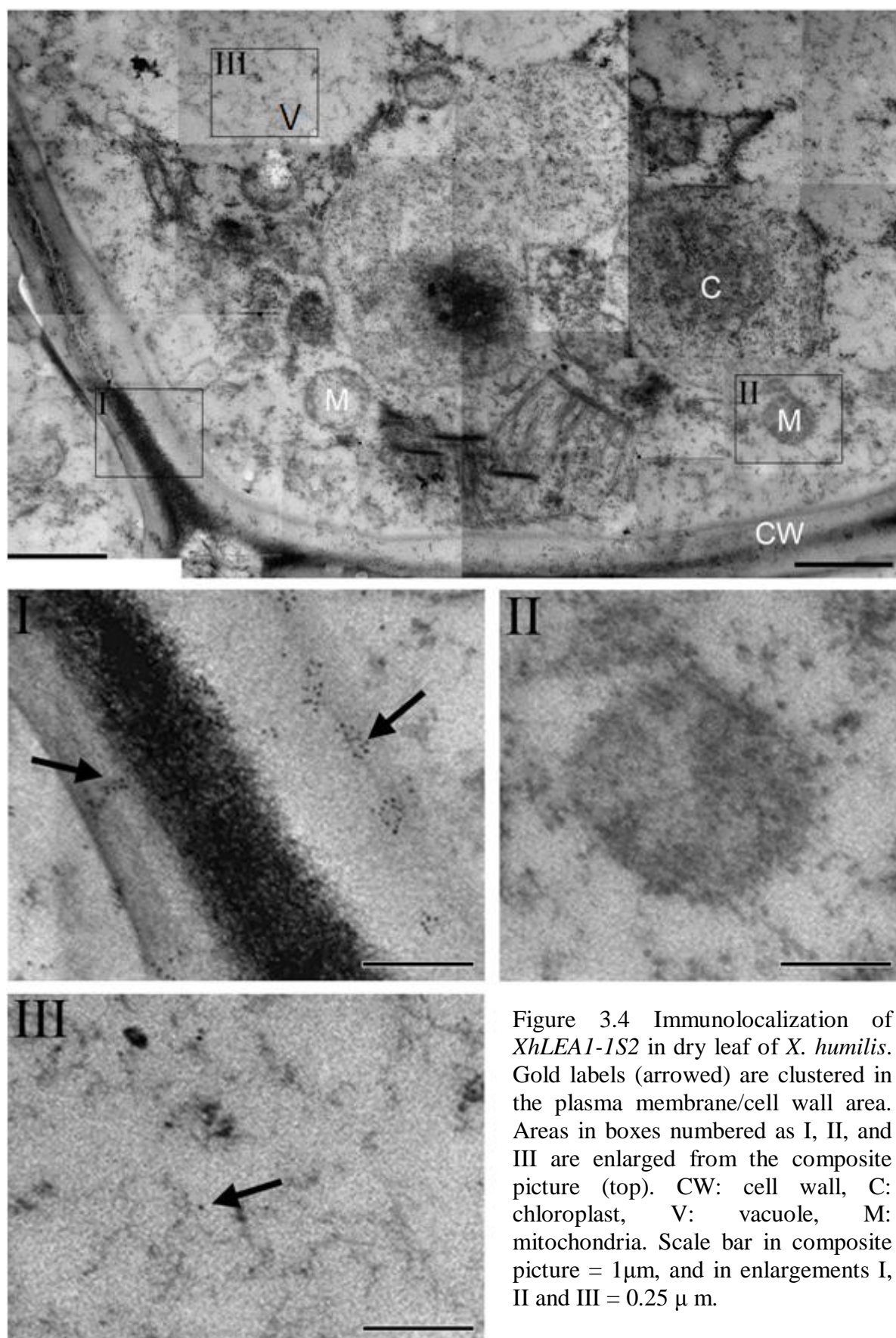


Figure 3.1 Immunolocalization of *XhLEA1-4S1* in hydrated leaf of *X. humilis*. Gold labels (arrowed) are randomly distributed in the different cellular compartments. Areas in boxes numbered as I, II, and III are enlarged from the composite picture (top). CW: cell wall, C: chloroplast, V: vacuole, M: mitochondria. Scale bar in composite picture = 1 μ m, and in enlargements I, II and III = 0.25 μ m.







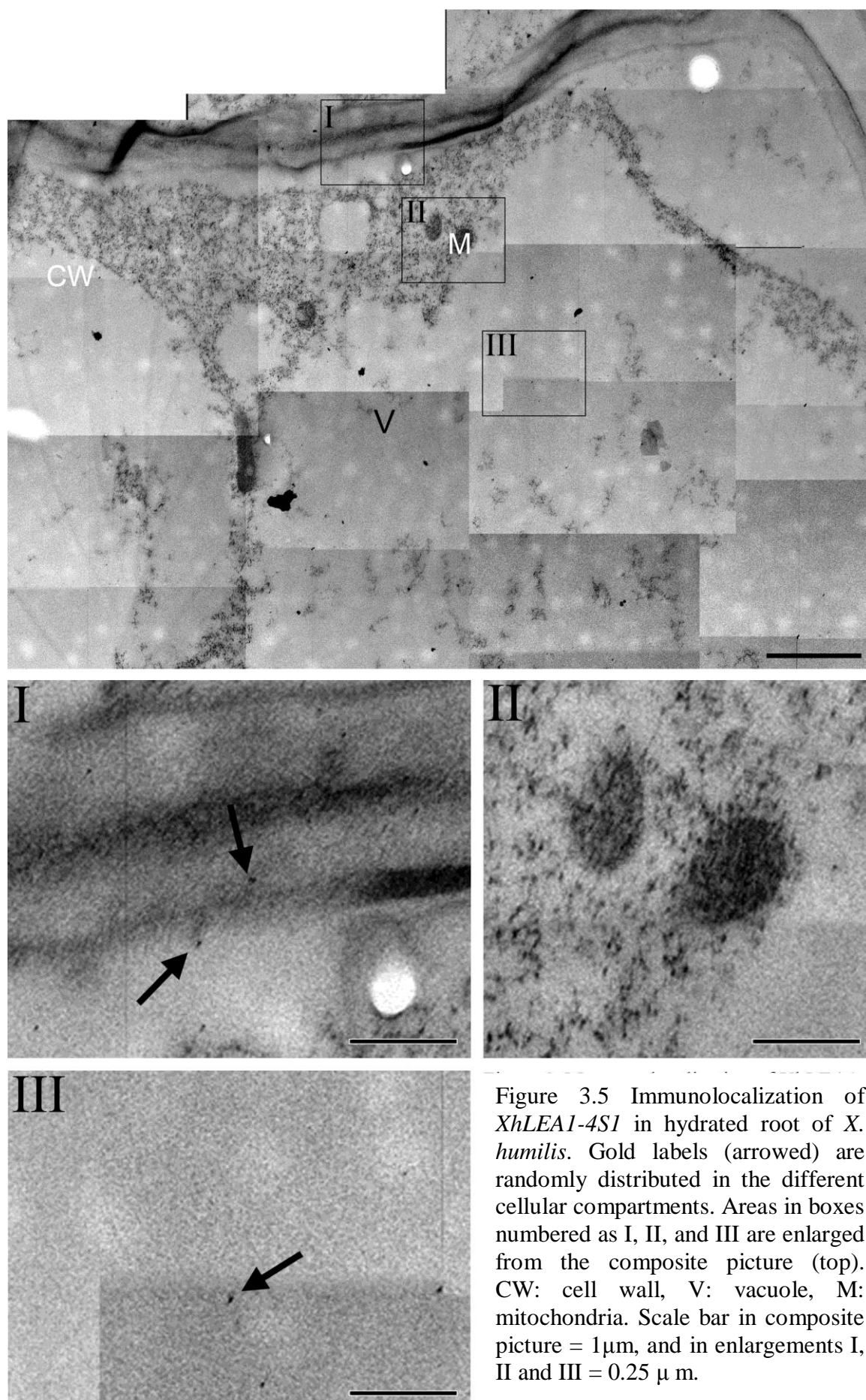
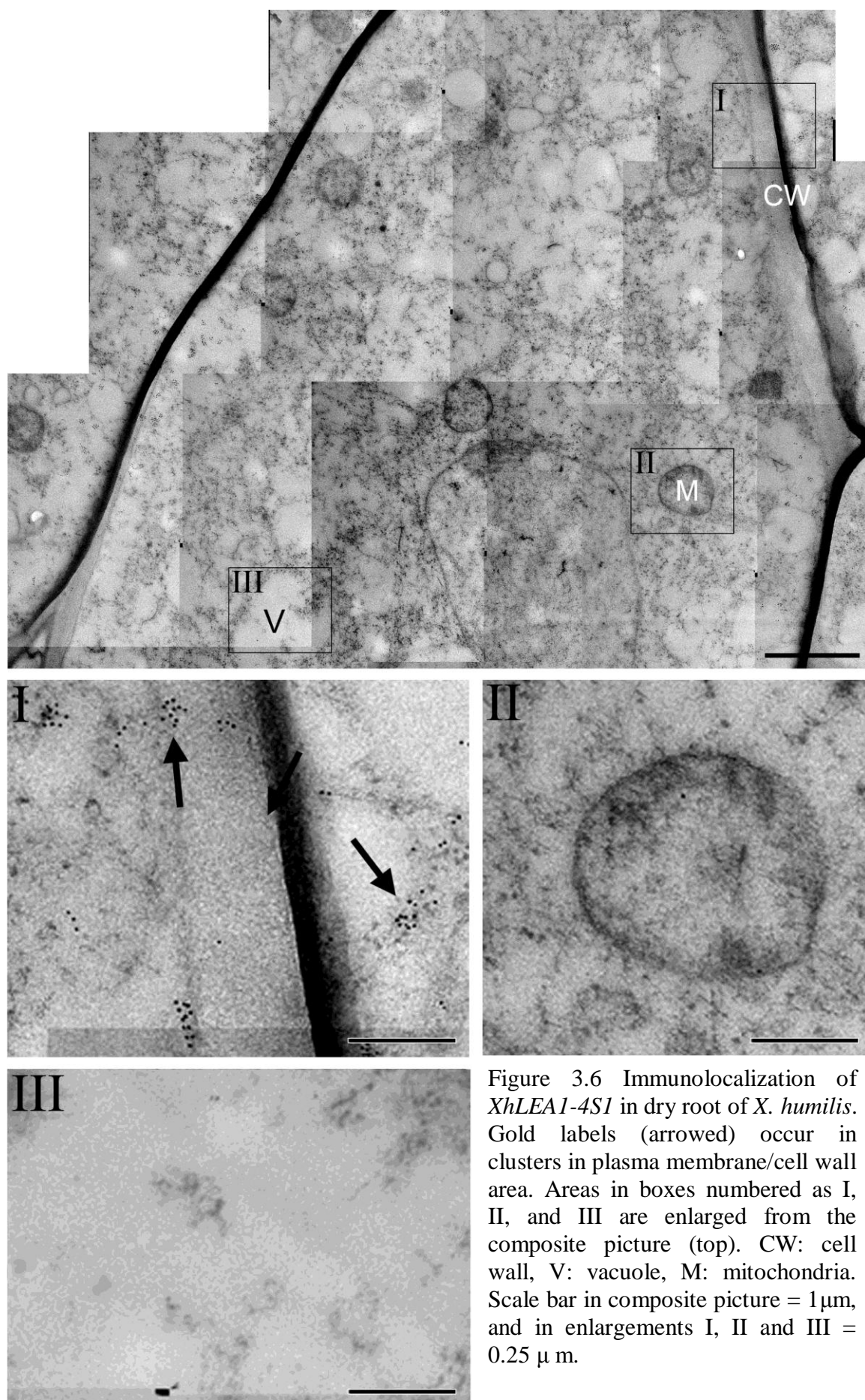
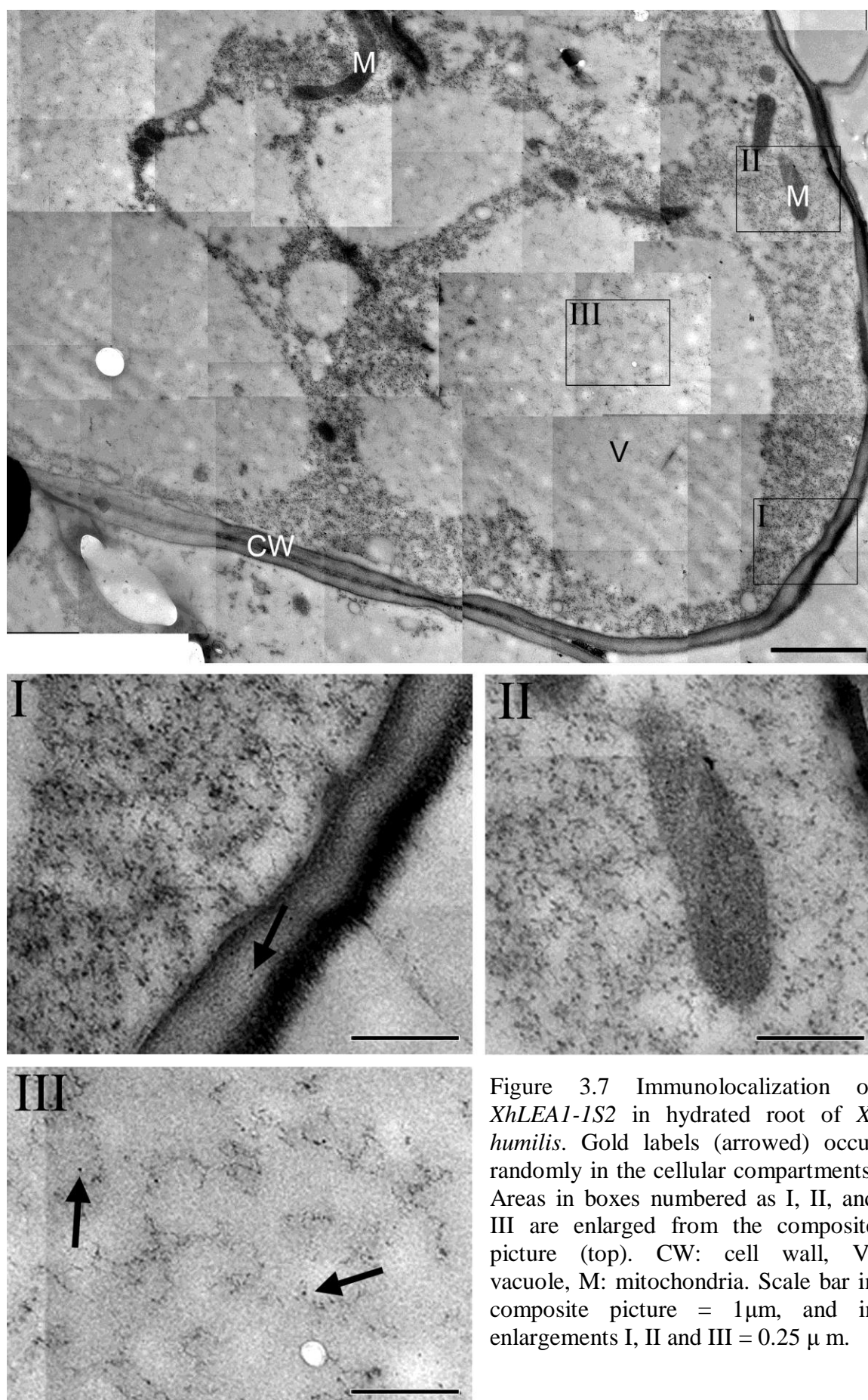
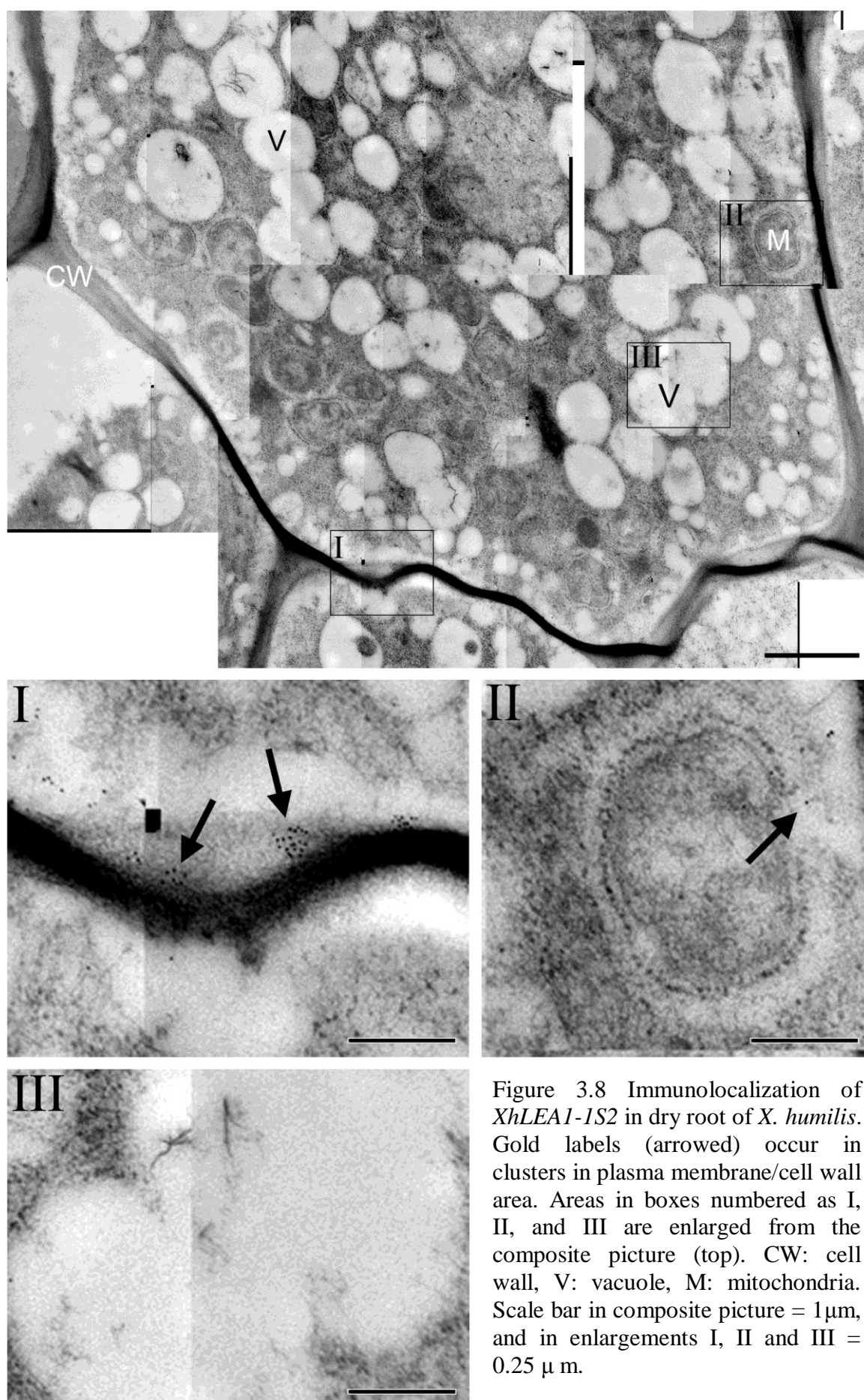


Figure 3.5 Immunolocalization of *XhLEA1-4S1* in hydrated root of *X. humilis*. Gold labels (arrowed) are randomly distributed in the different cellular compartments. Areas in boxes numbered as I, II, and III are enlarged from the composite picture (top). CW: cell wall, V: vacuole, M: mitochondria. Scale bar in composite picture = $1\mu\text{m}$, and in enlargements I, II and III = $0.25\mu\text{m}$.







3.3.1 Analysis of gold labelling by Contingency Table

Immunogold label distribution in cellular compartments of hydrated and dry samples of leaves was analysed using contingency table analysis as described earlier. The total χ^2 values for *XhLEA1-4S1* and *XhLEA1-1S2* proteins in leaves were 61.90 and 94.63 respectively (Tables 3.2 and 3.3). With degrees of freedom of 4 {(5 – 1 compartments) x (2 – 1 groups)}, the corresponding p-values was <0.0001. Hence the null hypothesis, i.e., there is no difference in gold labelling between hydrated and dry samples was rejected. This means expression levels of *XhLEA1-4S1* and *XhLEA1-1S2* in hydrated and dry states of the tissues studied was different.

Table 3.2 Analysis of gold label distribution in hydrated and dry *X. humilis* leaf cells using anti-*XhLEA1-4S1* primary antibody. Values represent observed (expected) counts of 15 nm gold labels. For a total χ^2 value of 61.90 and 4 degrees of freedom, $p < 0.0001$. The labelling distributions in hydrated and dry samples are significantly different and the partial χ^2 value indicates preferential labelling for cell wall + plasma membrane compartment in dry leaf.

Cell compartments	Observed (Expected) numbers of gold labels		Row sum	χ^2 value	
	hyd leaf	dry leaf		hyd	dry
Vacuole	213, (188.7056)	303, (327.2944)	516	3.127704	1.803317
Chloroplast	103, (82.65015)	123, (143.3499)	226	5.010474	2.888852
Mitochondria	51, (47.54212)	79, (82.45788)	130	0.251502	0.145006
Cell wall + plasma membrane	155, (232.9564)	482, (404.0436)	637	26.08728	15.04095
Cytosol + rest of the cell	216, (186.1457)	293, (322.8543)	509	4.788077	2.760626
Column sums	738	1280	2018	61.90379	

Table 3.3 Analysis of gold label distribution in hydrated and dry *X. humilis* leaf cells using anti-*XhLEA1-IS2* primary antibody. Values represent observed (expected) counts of 15 nm gold labels. For a total χ^2 value of 94.63 and 4 degrees of freedom, $p < 0.0001$. The labelling distributions in hydrated and dry cells are significantly different, and the partial χ^2 value indicates preferential labelling for cell wall + plasma membrane compartment in dry leaf.

Cell compartments	Observed (Expected) numbers of gold labels		Row sums	χ^2 values	
	hyd leaf	dry leaf		hyd	dry
Vacuole	119 (84.59823)	93 (127.4018)	212	13.98944	9.289366
Chloroplast	149 (119.3154)	150 (179.6846)	299	7.385244	4.904003
Mitochondria	33 (38.30863)	63 (57.69137)	96	0.735646	0.488489
Cell wall + plasma membrane	113 (191.5432)	367 (288.4568)	480	32.20699	21.38632
Cytosol + rest of cell	173 (153.2345)	211 (230.7655)	384	2.549514	1.692947
Column sums	587	884	1471	94.62795	

Observed and expected gold label counts and the corresponding partial χ^2 values were compared to determine preferential labelling of cellular compartments. Partial χ^2 values for cell wall-plasma membrane compartment of hydrated and dry leaf for both *XhLEA1-4S1* (26.09 and 15.04) and *XhLEA1-IS2* (32.21 and 21.39) proteins indicate that this compartment was preferentially labelled for both proteins (Tables 3.2 and 3.3). Each of the partial χ^2 values for this compartment contributed more than 10 % of the total χ^2 values and hence satisfies the criteria used for preferential labelling (Griffiths *et al.*, 2001; Mayhew *et al.*, 2004; Mayhew 2005). The observed gold label counts in cell wall-plasma membrane compartment for the *XhLEA1-IS2* in hydrated sample was less than expected and hence only the dry samples showed preferential labelling.

Similarly, partial χ^2 value corresponding to vacuoles of hydrated leaf sample (13.99) for *XhLEA1-IS2* protein indicate a significant difference between observed and expected number of gold label counts. The observed gold particle count is significantly larger than expected (119, 84.60) indicating preferential labelling of vacuoles of hydrated samples.

However, whereas labelling in the cell wall-plasma membrane area was in clusters of >5 gold particles, the labelling in vacuoles was random.

Gold label distribution in roots was analysed in the same way as for leaves. However, plastids could not be identified for certain; hence only four compartments were compared. The total χ^2 values for *XhLEA1-4S1* and *XhLEA1-IS2* proteins were 59.63 and 142.57 respectively (Table 3.4 and 3.5). With degrees of freedom of 3, $\{(4 - 1 \text{ compartments}) \times (2 - 1 \text{ groups})\}$, the corresponding p-values was <0.0001. Hence the alternative null hypothesis, i.e., there is no difference in gold labelling between hydrated and dry roots was rejected, and that expression of *XhLEA1-4S1* and *XhLEA1-IS2* in hydrated and dry roots was different.

As in leaves, the cell wall-plasma membrane compartment was preferentially labelled for *XhLEA1-IS2* (partial χ^2 values of 25.39856) in dry roots (Table 3.5). However, despite the high gold label counts (877) from the dry sample, and the appearance of clusters of labels, the partial χ^2 values for *XhLEA1-4S1* protein (1.102907) did not indicate preferential labelling of cell wall-plasma membrane compartment (Table 3.4). The substantial contribution to the total χ^2 value for *XhLEA1-4S1* in roots came from the vacuole compartment. Furthermore, the total gold count for *XhLEA1-4S1* in the dry roots was about 7 times more than that of hydrated state (Table 3.4). For *XhLEA1-IS2*, the total gold label counts in the hydrated and dry states did not change significantly (Table 3.5).

Observed gold label counts from vacuoles of hydrated roots were significantly higher than the expected values resulting in partial χ^2 values of 42.93235 for *XhLEA1-4S1*, and 41.49173 for *XhLEA1-IS2* respectively (Table 3.4 and Table 3.5). These results indicate that vacuoles of fully hydrated roots were preferentially labelled for both *XhLEA1-4S1* and

XhLEA1-1S2 proteins. However, the gold label distribution in vacuoles like in the rest of the compartments except cell wall-plasma membrane appeared to be random and non-specific.

Table 3.4 Analysis of gold distribution in hydrated and dry *X. humilis* root cells using anti-*XhLEA1-4S1* primary antibody. Values represent observed (expected) counts of 15 nm gold labels. For a total χ^2 value of 59.63 and 3 degrees of freedom, $p < 0.0001$. The labelling distributions are significantly different and the partial χ^2 value indicates preferential labelling in vacuoles of hydrated root cells.

Cell compartments	Observed (Expected) numbers of gold labels		Row sums	χ^2 values	
	hyd root	dry root		hyd root	dry root
Vacuole	38 (13.72532)	69 (93.27468)	107	42.93235	6.317471
Mitochondria	6 (3.84822)	24 (26.15178)	30	1.203194	0.177049
cell wall +Plasma membrane	94 (124.5541)	877 (846.4459)	971	7.495145	1.102907
Cytoplasm + rest of cell	53 (48.8724)	328 (332.1276)	381	0.348604	0.051297
Column sums	191	1298	1489	59.62801	

Table 3.5 Analysis of gold label distribution in hydrated and dry *X. humilis* root cells using anti-*XhLEA1-1S2* primary antibodies. Values represent observed (expected) counts of 15 nm gold labels. For a total χ^2 value of 142.57 and 3 degrees of freedom, $p < 0.0001$. The labelling distributions are significantly different and the partial χ^2 value indicates preferential labelling in vacuoles of hydrated roots and cell wall + plasma membrane compartment of dry roots.

Cell compartments	Observed (Expected) numbers of gold labels		Row sums	χ^2 values	
	hyd root	dry root		hyd root	dry root
Vacuole	114 (62.9096)	17(68.0904)	131	41.49173	38.33475
Mitochondria	36 (28.33333)	23 (30.66667)	59	2.07451	1.916667
Plasma membrane + cell wall	141 (218.5028)	314 (236.4972)	455	27.49021	25.39856
Cytoplasm + rest of the cell	134 (115.2542)	106 (124.7458)	240	3.048943	2.816958
Column sums	425	460	885	142.5723	

The *X. humilis* seeds used for immunolocalization study were at different levels of maturation. This was apparent from the variation in protein storage vacuole morphology in the different seeds. It is known that in seeds accumulating storage proteins, only partially filled vacuoles are seen at the start of reserve accumulation and that these bodies fill as seed maturation (and desiccation tolerance) occurs. Thus, the seeds can be aged accordingly.

Filled protein storage vacuoles (PSVs), where identified as dark bodies surrounded by narrow electron transparent areas (Figure 3.9 and 3.10), whereas in immature and partially mature seeds these were seen as spherically shaped bodies with light dark areas dispersed and also surrounded by electron transparent areas. Furthermore, in all the cells examined, lipid bodies were not clearly distinguishable.

Immuno-gold labelling of *X. humilis* seeds showed different results from those of leaves and roots. Gold labels for both *XhLEA1-4S1* and *XhLEA1-1S2* proteins were found associated with PSVs rather than cell wall or plasma membrane. These labels were seen within or at the close vicinity of these PSVs occurring in clusters (Figure 3.9 and 3.10). The clustering pattern of gold labels in seeds was similar as described earlier for cell wall-plasma membrane compartment of leaves and roots. Identification of gold labels in mature PSVs was difficult due to the dark colour of these structures but were able to be counted. Gold labels observed in the rest of the cellular structures were few in number and randomly scattered and hence regarded as non-specific.

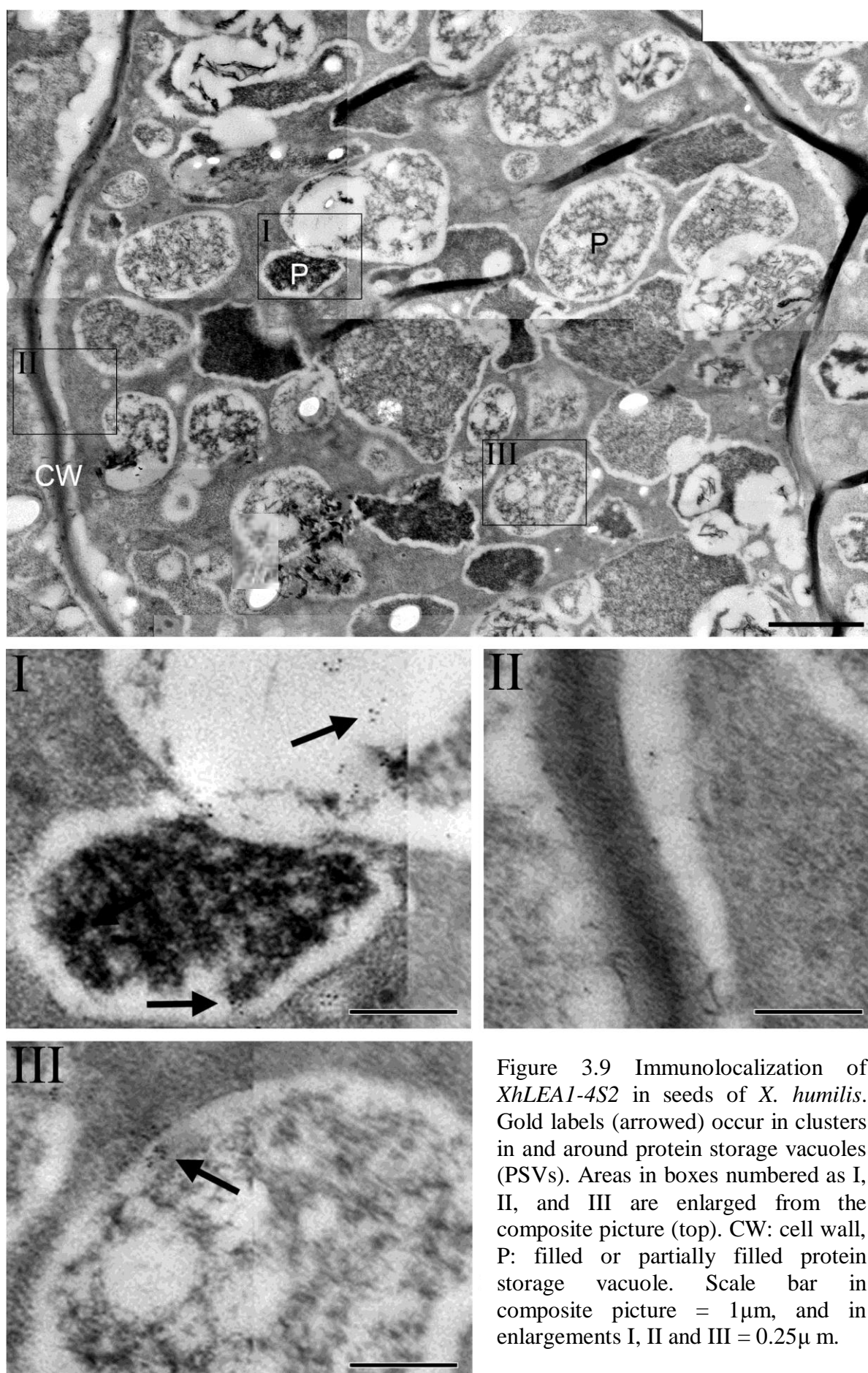


Figure 3.9 Immunolocalization of *XhLEA1-4S2* in seeds of *X. humilis*. Gold labels (arrowed) occur in clusters in and around protein storage vacuoles (PSVs). Areas in boxes numbered as I, II, and III are enlarged from the composite picture (top). CW: cell wall, P: filled or partially filled protein storage vacuole. Scale bar in composite picture = $1\mu\text{m}$, and in enlargements I, II and III = $0.25\mu\text{m}$.

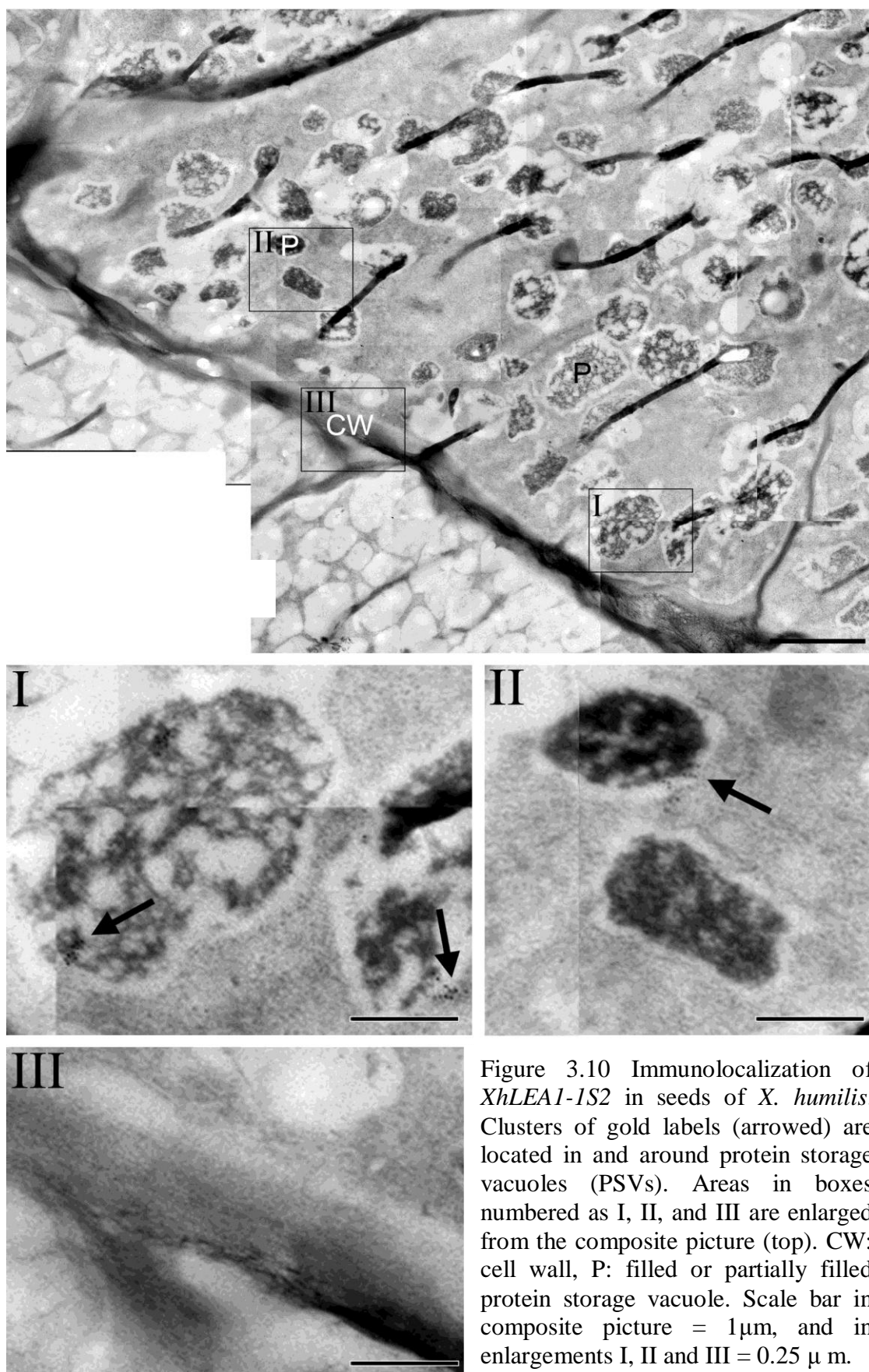


Figure 3.10 Immunolocalization of *XhLEA1-1S2* in seeds of *X. humilis*. Clusters of gold labels (arrowed) are located in and around protein storage vacuoles (PSVs). Areas in boxes numbered as I, II, and III are enlarged from the composite picture (top). CW: cell wall, P: filled or partially filled protein storage vacuole. Scale bar in composite picture = 1 μm, and in enlargements I, II and III = 0.25 μm.

3.3.2 Analysis of gold labelling by Pair Correlation Function (PCF)

As described earlier, the ContingencyTable analysis does not describe pattern of labelling, but it indicates preferential labelling of cellular compartments. Therefore, the pattern of gold-label distribution observed in the cell wall-plasma membrane compartment of leaves and roots as well as PSVs of seeds of *X. humilis* was further investigated using a pair correlation function (PCF) $g(r)$ as described by Gathercole *et al.* (2000) and D'Amico and Skarmoutsou (2008). As shown in Table 3.6, the population $g(r)$ values for both *XhLEA1-4S1* and *XhLEA1-1S2* in dry leaf and dry root samples as well as seeds around the 10 randomly selected reference points on the specific areas studied were greater than unity ($g(r) > 1$). On the other hand, $g(r)$ values for fully hydrated leaf and root samples for similar areas were between 0.6 and 0.8, a value indicative of mixtures of regular and random gold label distribution.

Table 3.6 Analysis of gold label distribution patterns for *XhLEA1-4S1* and *XhLEA1-1S2* proteins in the cell wall-plasma membrane compartment of fully hydrated and dry *X. humilis* leaf and root tissues by Pair correlation function, $g(r)$. Gold labels were counted from concentric circles of radius r and $r+\Delta r$ around 10 different reference gold particles for each sample and used for $g(r)$ calculation. $g(r) > 1$ indicates clustering.

r (nm)	Pair Correlation Function (PCF, $g(r)$)							
	<i>XhLEA1-4S1</i>				<i>XhLEA1-1S2</i>			
	hyd leaf	dry leaf	hyd root	dry root	hyd leaf	dry leaf	hyd root	dry root
50	2.66	6.83	0.56	9.70	1.88	6.84	1.22	5.96
100	0.74	1.89	1.68	2.95	0	2.13	0.77	2.60
150	0.69	0.73	0.56	0.98	0.31	1.18	0.20	1.29
200	0.30	0.58	0.70	0.29	0	0.66	0.46	0.63
250	0.71	0.66	1.23	0.16	0.56	0.43	1.04	0.32
300	0.49	0.32	0.28	0.16	1.72	0.21	0.66	0.15
350	0.30	0.24	0.40	0.18	0.27	0.22	0.66	0.23
400	0.63	0.18	0.35	0.15	0.71	0.15	0.27	0.29
Population $g(r)$	0.81	1.43	0.72	1.82	0.68	1.48	0.66	1.43

Table 3.7 Analysis of gold label distribution patterns for *XhLEA1-4S1* and *XhLEA1-1S2* proteins in protein storage vacuoles (PSVs) of *X. humilis* by Pair correlation function, $g(r)$. Gold labels were counted from concentric circles of radius r and $r+\Delta r$ around 10 different reference gold particles for each sample and used for $g(r)$ calculation. $g(r)>1$ indicates clustering.

r (nm)	Pair Correlation Function (PCF, $g(r)$)	
	<i>XhLEA1-4S1</i>	<i>XhLEA1-1S2</i>
0	1	1
50	7.8544031	8.545866
100	3.1842175	2.756731
150	0.5660831	0.964856
200	0.0530703	0.344591
250	0.0849125	0.330808
300	0.2122812	0.206755
350	0.3335847	0.295364
400	0.2653515	0.155066
Population $g(r)$	1.569238	1.700005

When $g(r)$ values of gold particles at a given distance from the reference particles were plotted against interparticles separation, r , a typical curve corresponding to clustering distribution was observed (Figure 3.11). The highest peak of the curves for dry leaf and dry root as well as for the PSVs of seeds occur at around a distance r of about 50 nm. This means most gold labels are found close to each other with an average interparticles distance of 50 nm for dry samples. For *XhLEA1-4S1* and *XhLEA1-1S2* proteins from hydrated leaf and hydrated roots, the population $g(r)$ was below 1 indicating that gold particles were randomly scattered across the study area.

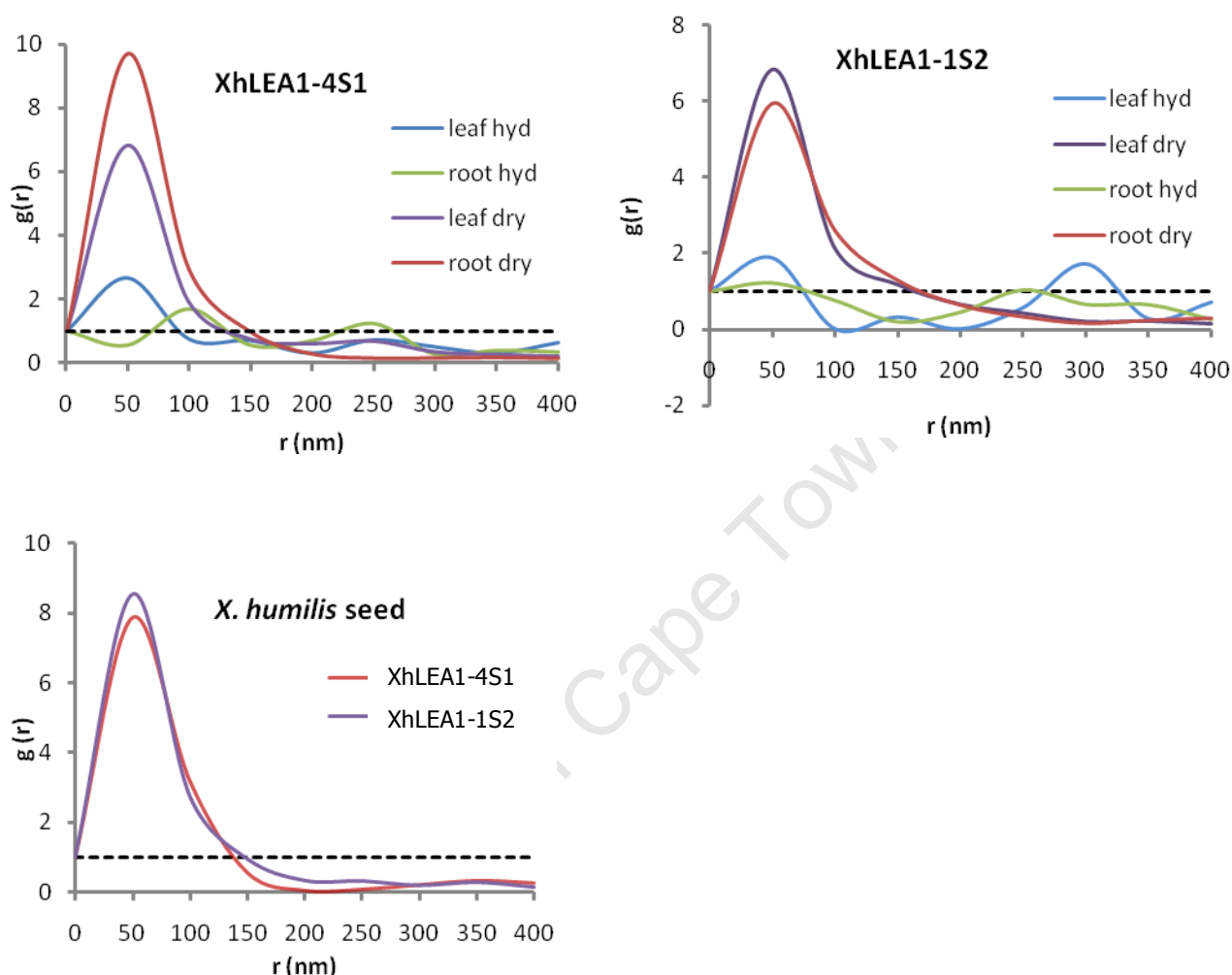


Figure 3.11 Test for clustering of gold labels for *XhLEA1-4S1* (top left) and *XhLEA1-1S2* (top right) proteins in cell wall-plasma membrane compartments of hydrated and dry leaf and root samples and PSVs of seeds (bottom) of *X. humilis* by Pair Correlation Function, $g(r)$. Immunogold labels in these structures were counted from selected areas of $642.86 \mu\text{m}^2$ by constructing concentric circles of 50 nm width (Δr) around 10 reference gold particle from which mean $g(r)$ was calculated. Mean $g(r)$ of samples was plotted as a function of interparticle separation, r in nm. $g(r) > 1$ indicates clustering (blue and green lines for hydrated, and red and pink lines for dry leaf and root samples respectively) and $g(r) < 1$ indicates regular distribution. $g(r) = 1$ (broken line) indicates complete random distribution.

3.4 Discussion

Paraformaldehyde fixation was chosen in the current study as it allows better penetration of antibodies to target antigens and also results in less background compared to glutaraldehyde and osmium tetroxide fixatives (Scarano *et al.*, 2003). In addition to higher background staining, osmium fixation is not compatible with antibody-based antigen localization studies. Although penetration is good, paraformaldehyde, fixes tissues slowly (Pow *et al.*, 1995; Kiernan 2000 Marttila and Santén, 2007). This slow fixation can cause osmotic imbalance which can lead to structural crumple before the cellular compartments are fixed. This might have caused the detachment of cytoplasm from the cell wall observed in some of the cells of the tissues studied.

The absence of gold labelling in samples where no primary antibody was added showed that there was no background labelling and that all the labels observed in the rest of the samples reflected specific or non-specific affinity of the primary antibodies to those XhLEA proteins in the different cellular compartments. From observation of electron micrographs of leaf and root cells, it was apparent that cell wall-plasma membrane area was heavily labelled by clusters of gold particles. This preferential labelling was confirmed by contingency table analysis which showed that *XhLEA1-4S1* and *XhLEA1-1S2* were localized in the cell wall-plasma membrane area.

Although some labelling was observed in some compartments of hydrated tissues for both *XhLEA1-4S1* and *XhLEA1-1S2* proteins, this labelling was discounted as non-specific on the basis that expression of the respective proteins was not detected earlier by western blotting (Chapter 2). The PCF data of hydrated tissues (Table 3.6) also showed some individual points, $g(r)$, where values were over 1.0. However, the population data remained to be below

1.0 indicating that the observed labelling in the hydrated tissues was once again non-specific.

Contingency table analysis results showed preferentially labelled compartments only and did not indicate the pattern of gold label distribution. PCF, on the other hand, has the strength in describing the pattern of gold label distribution. The latter identified the pattern of gold labelling within compartments as random or clustered. Clustering of labels around specific areas and structures was related to the affinity of the antibodies to the proteins in them as opposed to the rest of the intracellular spaces. Hence, the results obtained indicated the presence of *XhLEA1-4S1* and *XhLEA1-1S2* associated with the cell wall-plasma membrane of dry tissues of leaves and roots and the PSVs of seeds of *X. humilis*.

Internal organization of the *X. humilis* seeds studied showed PSVs at different levels of filling indicating that the seeds were not all completely mature. In seeds, PSVs are known to store storage and defence proteins (Neuhaus and Rogers, 1998). Immunolabelling results of this study have shown that *XhLEA1-4S1* and *XhLEA1-1S2* proteins occur in or around these PSVs in *X. humilis* seeds. These might be an indication that these proteins play a role related to PSV function, or maybe they are stored there until required. In addition to their role in desiccation tolerance, Group 1 LEA proteins have been proposed to be nutritional storage proteins at least in one earlier report (Swire-Clark and Marcotte, 1999). A barley (*Hordeum vulgare* L) Group 3 LEA protein is also reported to be abundant in PSVs (Marttila *et al.*, 1996).

The localization results of *XhLEA1-4S1* and *XhLEA1-1S2* proteins in *X. humilis*, dry leaves and dry roots and not in the hydrated state were indications that these proteins are desiccation linked and confirmed results obtained earlier in Western blot expression study

(Chapter 2). Furthermore, the location of *XhLEA1-4S1* and *XhLEA1-1S2* proteins in the cell wall-plasma membrane area, at least in the vegetative tissues studied, indicate that these proteins may be involved in stabilizing membrane or membrane associated structures. The hydrophilic nature of LEA proteins was reported to enable them to form hydrogen bonding with other molecules (Rajesh and Manickam, 2006; Battaglia *et al.*, 2008). Therefore, it is possible that these XhLEAs form extensive hydrogen bonding with membrane structures preventing structural collapse as the result of water loss. However, further study is required to elucidate these interactions.

The absence of *XhLEA1-4S1* and *XhLEA1-1S2* in cell wall-plasma membrane area of seeds may be related to the developmental stage of the seeds described above. It is possible that the seeds did not reach desiccation tolerant stage when the LEA proteins are required for membrane stabilization. This could be investigated by future study using seeds at different levels of maturation. It is also worth mentioning that the fixation method used in the current study was not the best of its kind in terms of organelle definition. Better ultrastructural resolution could be obtained by optimization of fixation method. This is the first report that expression and localization of *XhLEA1-4S1* and *XhLEA1-1S2* protein in *X. humilis* leaves, roots and seeds have been described, and the results could be used as a basis for future studies in establishing the exact function of these XhLEA proteins in desiccation tolerance.

Chapter 4

Functional and structural characterization of *XhLEA1-4S1* and *XhLEA1-1S2*

4.1 Introduction

Several putative functions have been attributed to the different LEA proteins in relation to water loss from seeds and vegetative tissues of desiccation tolerant plants and some anhydrobiotic organisms (reviewed in Chapter 1). However, most of these proposed functions have not been supported by functional studies. It is also not known if some of the proposed functions are specific to specific groups of LEAs or if all LEAs have similar mechanisms of action against desiccation stress. LEA proteins that belong to the same group have been reported to be induced under different stress conditions (Dalal *et al.*, 2009). Therefore, functional and structural characterization of individual LEA proteins is important in establishing the role of these proteins in desiccation tolerance.

Some biochemical assays have been developed to investigate the stabilization role of LEAs on other proteins under desiccation stress. Goyal *et al.* (2005) have used an aggregation assay to demonstrate the role of the Wheat *Em*, a Group 1 LEA, and *AavLEA1*, a Group 3 LEA protein from a nematode (*Aphelenchus avenae*). In the presence of these LEAs, aggregation of lactate dehydrogenase (LDH) and Citrate synthase enzymes was prevented; and enzyme activity maintained. In a similar study, Chakrabortee *et al.* (2007) investigated if *AavLEA1* protein could prevent desiccation-induced aggregation of a complex protein mixture *in vitro* and *in vivo*. In the latter assay, water-soluble proteomes isolated from the human cell line T-REx293 (derived from the embryonic kidney cell line *HEK293*) and the nematode *A. avenae* were subjected to *in vitro* desiccation in the presence of *AavLEA1* and no aggregation was observed. They have also showed *in vivo* anti-aggregation role for *AavLEA1* by expressing it in a human cell line (*T-REx293-LEA15*) and measuring the

aggregation of a cytoplasmic protein (*EGFP-HDQ74*) on dehydration in the presence or absence of *AavLEA1*.

In a similar study, Rinne *et al.* (1999) reported that an abscisic acid-responsive dehydrin, Group 2 LEA, maintained the activity of alpha-amylase enzyme under conditions of low water activity. These researchers used 20% polyethylene glycol (PEG) to simulate a dehydrating condition and the enzyme activity was measured in the presence or absence of the dehydrin. Enzyme activity was maintained in the presence of the dehydrin when compared with a sample from which the dehydrin was removed by immunoprecipitation. This result showed that hydrophilic LEA proteins can act as hydration buffers by pooling water through their hydrophilic portions providing the intermolecular motions necessary for maintaining enzyme activity.

Knowledge of the structure of a biological molecule is an important part of the characterization as it gives clues to function. The traditional approach to study a protein involves crystallization. However, the LEA proteins studied so far are not structured in solution and hence crystallization approaches to study their structure are not useful. It is also not known if LEA proteins require structure to have a biological function. In recent years efforts have been made to determine what structure LEA proteins would assume in dry state using other alternative approaches such as Circular Dichroism (CD).

CD spectroscopy measures the differences in the absorption of left-handed polarized light versus right-handed polarized light which arises due to structural asymmetry. Far-UV CD spectra (190-250 nm) requires 20 – 200 μ l of solution containing 1 mg/ml to 50 μ g/ml protein in any buffer which does not have a high absorbance in this region of spectrum. The three main protein structures, α -helices, β -sheets, and random coils, display characteristic spectra at wavelengths of 190 to 250 nm and the transition from one structure to another can

be monitored by CD. The transition from one structural confirmation to another can be induced by the use of structure inducers such as trifluoroethylene (TFE), glycerol and SDS (Mouillon *et al.*, 2008; Shih *et al.*, 2010) or by denaturant substances such as guanidinium chloride (GdmCl, Soulages *et al.*, 2003). Structural studies of LEA proteins using these methods have shown that these proteins are unstructured in solution and display a spectrum similar to random coil (Soulages *et al.*, 2002; Gilles *et al.*, 2007; Mouillon *et al.*, 2008; Shih *et al.*, 2010a).

There are also bioinformatics tools such as PONDR (Prediction of Naturally Disordered Regions) by which the disorder of a LEA protein can be predicted. PONDR, is a web based program available at the developer's website, <http://www.pondr.com>. This program has been used successfully to predict protein disorder in many reports (Obradovic *et al.*, 2005; Peng *et al.*, 2006; Fuxreiter *et al.*, 2007). Both CD and PONDR were used in this study to investigate the structure of *XhLEA1-4S1* and *XhLEA1-1S2* proteins from *X. humilis*.

The objectives of the study in this chapter were (1) to investigate the in-solution structure of *XhLEA1-4S1* and *XhLEA1-1S2* proteins and (2) determine if these proteins have a stabilizing/protective role on other proteins under drying conditions. Protein structure was investigated using Circular dichroism (CD) spectroscopy and PONDR algorithm. The assays conducted were aggregation and enzyme activity of citrate synthase (CS) enzyme dried in the presence or absence of either of these XhLEAs.

4.2 Methods and materials

4.2.1 Construction of pET-21a(+) expression plasmids

Primer sets for cloning of *XhLEA1-4S1* and *XhLEA1-1S2* in pET-21a(+) vector (Table 4.1) were designed as described for pGEX-3X vector in Chapter 2. These primers were put into *XhLEA1-4S1* and *XhLEA1-1S2* cDNA sequences through PCR using previously constructed pGEM- plasmids of these clones (Chapter 2) as templates (PCR: 30 cycles of 94 °C, 62 °C, 72 °C, and a final extension at 72 °C for 5min). PCR amplified DNA bands of *XhLEA1-4S1* and *XhLEA1-1S2* were then gel purified, sequentially restriction digested and ligated in a 3:1 insert:vector ratio into pET-21a(+) vector (Vector map, Appendix F) which was also double-digested with the same enzymes. The ligation reaction was used to transform *DH5-alpha* cells, known to be easy to transform. The plasmids containing the respective LEA cDNAs were then isolated and used to transform *BL21(DE3)pLys* expression host cells. The proper placement of the newly constructed XhLEA cDNAs in pET-21a(+) vectors with respect to the vector's translation start codon was checked through sequencing.

Table 4.1 Primers used for cloning and sequencing of *XhLEA1-4S1* and *XhLEA1-1S2* in pET-21a(+) vector

Name	Sequence 5' --- 3'
1. <i>XhLEA1-4S1</i> Forward primer:	GTAGGATCCTTATGGCTTCCCATCAA
2. <i>XhLEA1-4S1</i> Reverse Primer:	GCTGAATTCACGAACTAAGAACGTCG
3. Xh0797 Forward Primer:	GTAGGATCCTATCAATCATGGCTTCTG
4. Xh0797 Reverse Primer	GCATGAATTCATCTCGTCGGAGATAGGT

NB. *Bam*HI and *Eco*RI sites are underlined and translation start codon is in bold.

4.2.2 Production of recombinant proteins

Expression induction, optimization and harvesting of bacterial lysates containing the T7 tagged recombinant proteins of *XhLEA1-4S1* and *XhLEA1-1S2* were as described for GST-

XhLEA fusion proteins in Chapter 2 except that the expression vector used here was pET-21a(+) (Novagen).

4.2.3 Fractionation of heat stable proteins

Total protein pellets of cultures from step 4.2.2 were re-suspended in ice-cold 1X T7•Tag Bind/Wash Buffer (1/20 of original culture volume) containing 1mM PMSF, and lysed by sonication in an ice/ethanol bath with a microtip attached to a sonicator (Misonix 3000), output power set at 4, 5 cycles of alternating 30 seconds ON and 60 seconds OFF to decrease foaming and cell debris were removed by centrifugation at 39,000 x g for 20 minutes.

To fractionate heat stable proteins only, a heat treatment step was introduced as previously reported for the isolation of heat stable proteins in a related work (Borovskii *et al.*, 2002; Pelah *et al.*, 1995; Rudiger *et al.*, 1995). Lysates containing *XhLEA1-4S1* and *XhLEA1-1S2* recombinant proteins were heated at 70, 80, 90 and 100 °C for 10 minutes. Denatured proteins were precipitated by centrifugation at 10,000 x g for 10 minutes. To check if the heat stable portion contained the XhLEA proteins, 15µg of protein in the supernatant was separated on a 12 % SDS-PAGE for analysis. Expression was scaled up and recombinant XhLEA proteins were affinity purified from samples heated at 80 °C for 10 minutes.

4.2.4 Isolation of T7-Tagged XhLEA recombinant proteins

The recombinant *XhLEA1-4S1* and *XhLEA1-1S2* proteins containing T7-tag were purified using T7•Tag Affinity Purification Kit as follows: Antibody bound agarose was transferred into two separate columns and equilibrated to room temperature, washed with ten column volumes (10 ml) of 1X T7•Tag Bind/Wash Buffer each. Partially purified *XhLEA1-4S1* or *XhLEA1-1S2* recombinant proteins from step 3.2.3 were then loaded onto these columns and

incubated for 30 minutes with gentle shaking. Unbound proteins were washed down with ten column volumes (10 ml) of 1X T7•Tag Bind/Wash Buffer. The recombinant proteins were eluted with 1X T7•Tag Elute Buffer and elute fractions were collected in tubes containing 1X T7•Tag Neutralization Buffer. Elute fractions for each protein were pooled, concentration determined (Bradford, 1976) and quality analysed on SDS-PAGE.

4.2.5 Structural studies on *XhLEA1-4S1* and *XhLEA1-1S2*

A. Prediction of structure from amino acid sequence

An online service called Predictor of Natural Disordered Regions (PONDR- VL-XT) (www.PONDR.com) was used to predict the overall structure of *XhLEA1-4S1* and *XhLEA1-1S2* amino acid sequences. FASTA formatted protein sequence of each of the LEA proteins was entered into the analysis tool. The pictorial and descriptive data output was analyzed for known LEA protein features and compared with published data on similar proteins. The extinction coefficients of the two proteins were also determined using the ExPASy proteomic server (<http://au.expasy.org/>).

B. In-solution structural investigation

To investigate the in-solution structure of *XhLEA1-4S1* and *XhLEA1-1S2*, CD measurements were taken from aqueous solution of these proteins in 10 mM PBS as well as in 50 % Trifluoroethylene (TFE) using JASCO J-810 Spectropolarimeter in the range of 185 to 260 nm wave length as described in the study of other LEA proteins (Soulages *et al.*, 2003; Shih *et al.*, 2004 and 2010; Mouillon *et al.*, 2006). Measurements of millidegrees sensitivity were taken from sample concentrations of 0.7 mg/ml using 0.1cm standard cuvette. The generated CD data was exported to Excel worksheet and used to plot molar ellipticity against wave length in nanometers (nm). The resulting CD spectra were compared with typical spectra of the three common forms of protein structures, α - helical, β -sheet and random coil. The percentage composition of these different protein structures was determined by the CD

deconvolution program, CDNN, (version 2.0.3.188) using molar ellipticity as a measurement unit.

4.2.6 LEA protein anti-aggregation role

To check if *XhLEA1-4S1* and *XhLEA1-1S2* proteins are involved in protecting other proteins from desiccation induced damage, CS enzyme (Cat. No. C 3260, Sigma) aggregation assay and activity tests were performed after 2 cycles of drying in the presence or absence of the LEA proteins as described by Goyal *et al.* (2005) and Chakrabortee *et al.* (2007). Briefly, LEA proteins and CS were dialysed against several changes of distilled water and the concentrations determined using a molar absorption coefficient of 1.78 for a 1 mg/ml solution at 280 nm at a 1 cm path length for CS and by Bradford (1976) method for the XhLEA proteins. Then 0.12 mg protein of *XhLEA1-4S1* and *XhLEA1-1S2*, each in a 200 μ l volume of solution, were first dried in a pilot experiment to establish that these proteins themselves do not aggregate when dried and re-hydrated. Proteins were dried either using a speedy-vacuum at room temperature, or a Dura-Stop™ microprocessor-controlled tray drier (FTS Systems) set at 25 °C, 2000 mT both for 2 hrs. Dried samples were dissolved in distilled water to the original volume and aggregation was measured by reading absorbance at 340 nm in a DU800 spectrophotometer (Beckman Coulter, High Wycombe, Bucks., U.K).

After establishing that the XhLEA proteins themselves do not aggregate, 0.12 mg of CS was dried in the presence or absence of *XhLEA1-4S1* or *XhLEA1-1S2* at CS:LEA molar ratios ranging from 1:0.5 to 1:10. Molar ratios that showed an apparent reduction of CS aggregation were repeated 3 times and results analyzed. The nematode *AavLEA1* protein known to prevent CS aggregation (Goyal *et al.*, 2005), provided by Dr Alan Tunnacliffe, University of Cambridge, was used as a positive control in all aggregation and enzyme activity experiments; BSA was also included as a negative control. Statistical significance of

results was determined by one-way ANOVA and Tukey post test using InStat3 (GraphPad Software).

4.2.7 LEA protein enzyme stabilization role

The effect of drying on CS activity in the presence or absence of the LEA protectants at CS:LEA molar ratios of 1:1 and 1:5 was also investigated as described by Goyal *et al.* (2005). Briefly, 2 μ l of CS/LEA (from the samples used in the aggregation experiments) was added to 1 ml of Tris/EDTA buffer with 100 μ M oxaloacetic acid, 100 μ M DTNB [5,5_-dithiobis-(2-nitrobenzoic acid)] and 150 μ M acetyl-CoA (disodium salt). Change in A412 was measured every 1.5 seconds for 1 min. Enzyme activity was expressed as a percentage of untreated CS, and all samples were assayed in triplicate. Statistical significance was determined as described in step 4.2.7 above.

4.3 Results

4.3.1 Cloning and production of XhLEA recombinant proteins

The GST-LEA fusion proteins produced in GST system as described in Chapter 2 could not be used for functional or structural analysis. This is because of the fact that the GST tail could not be removed and that any result of the functional assays could not be attributed to the LEAs for certain. The GST partner was bigger in size than the LEA proteins themselves and might mask any effect from the LEA proteins in functional assays. Therefore the *XhLEA1-4S1* and *XhLEA1-1S2* cDNA's were re-cloned in pET-21a(+) vector (Novagen) that adds only 11-amino acids at the N-terminus of the protein of interest.

The cloning procedure followed in constructing pGEX-XhLEA plasmids (Chapter 2) also worked for constructing the pET-XhLEA expression plasmids. PCR products obtained using the primers designed and pGEXT-*XhLEA1-4S1* and *XhLEA1-1S2* plasmids as templates

showed clear bands of DNA corresponding to the expected sizes (~ 0.5 and 0.45 kbs respectively; Figure 4.1). Restriction digestion, ligation of the XhLEA PCR products into pET-21(a+) vector as well as transformation of the expression host cells, *BL21 (DE3)pLysS*, with recombinant pET-XhLEA plasmids was also successful. Colony-PCR screening have shown the clones containing the cDNAs of either *XhLEA1-4S1* or *XhLEA1-1S2* in pET-21a(+) plasmid (Figures 4.2). Furthermore, no mutation was detected in the nucleotide sequences of the pET-XhLEA recombinant plasmids when checked by sequencing; and the open reading frame was found to be in line with the vectors translations start codon.

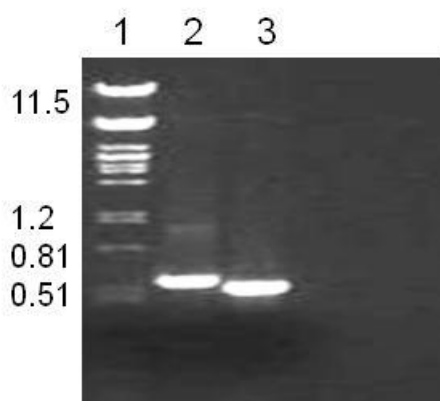


Figure 4.1 Insertion of *Bam*HI and *Eco*RI restriction sites into *XhLEA1-4S1* and *XhLEA1-1S2* cDNA sequences through PCR. Lanes 1 = λ *Pst*I molecular weight marker, 2 = *XhLEA1-4S1* cDNA, 3 = *XhLEA1-1S2* cDNA.

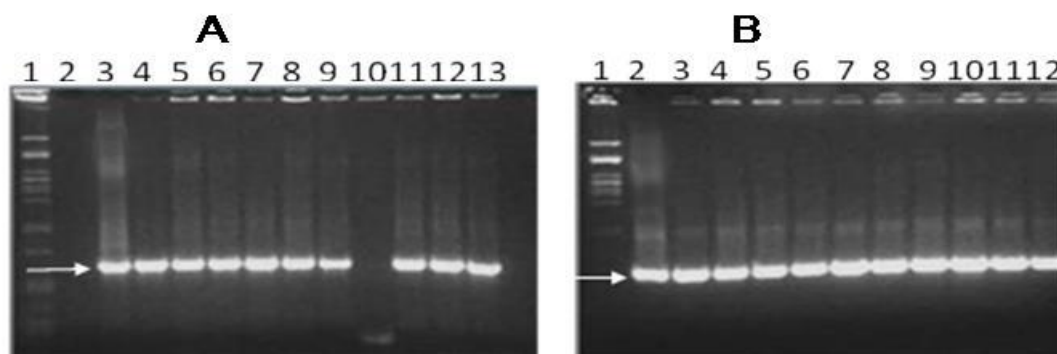


Figure 4.2 Colony-PCR screening of clones grown after transformation with recombinant pET-21a(+) plasmid containing either *XhLEA1-4S1* (A, lanes 3 -13) or *XhLEA1-1S2* (B, lanes 2 - 12) cDNAs using primers on Table 4.1. Lane 1 in both A&B is a molecular weight marker; lane 2 in A is a no template control.

pET-21a(+) expression vector adds 11 amino acids making the T7-Tag on the N-terminus of the recombinant proteins. These few amino acids were reported to have no known impact on the overall function of the tagged protein and in many published assays the biological function of such proteins were tested without removing the tag (Zheng *et al.*, 1999; Cartwright, 2001). The T7 Tag is also important in the purification of the recombinant protein with the T7-Tag affinity Purification Kit (Novagen)

XhLEA1-4S1 and *XhLEA1-1S2* recombinant proteins were produced by IPTG induction, 1mM at an OD₆₀₀ of 0.6 and harvested after a further incubation for 4 hours at 30 °C (Figure 4.3). The molecular weight of the expressed *XhLEA1-4S1* (+T7 Tag) recombinant protein on SDS gel appeared to be similar to the expected size (~17 kDa), but that of *XhLEA1-1S2* (~15 kDa) looked bigger than expected (Figure 4.3). However, western blot analysis confirmed that the observed bands corresponded to the respective recombinant proteins (Figure 4.4).

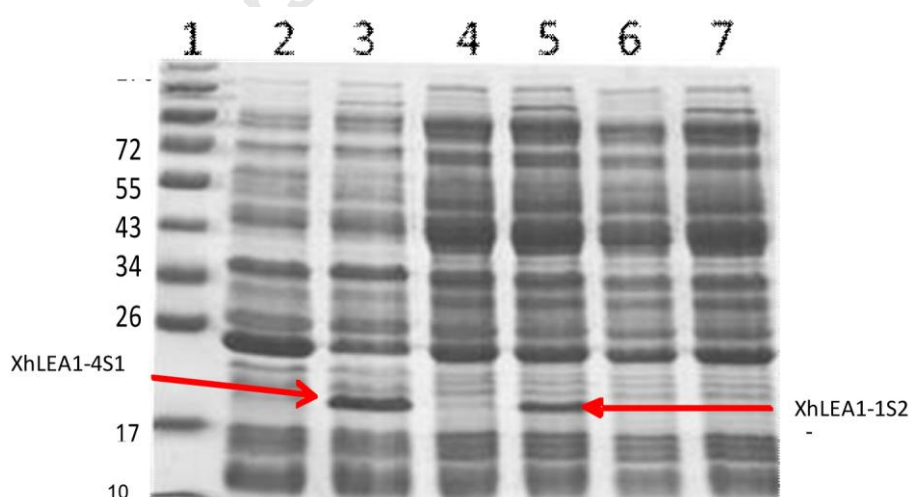


Figure 4.3 Expression of *XhLEA1-4S1* and *XhLEA1-1S2* recombinant proteins in pET-21a(+) expression vector. Lanes 1=molecular weight marker (Prestained Protein Ladder, #SM0671), 2= un-induced culture of cells containing pET 21+*XhLEA1-4S1* plasmid, 3=induced culture of cells containing pET 21+*XhLEA1-4S1* recombinant plasmid, 4= un-induced culture of cells containing pET 21+*XhLEA1-1S2* recombinant plasmid, 5= un-induced culture of cells containing pET 21+*XhLEA1-1S2* recombinant plasmid, 6 and 7=induced and un-induced culture of cells respectively containing empty vector. Expression was induced with the addition of 1mM IPTG.

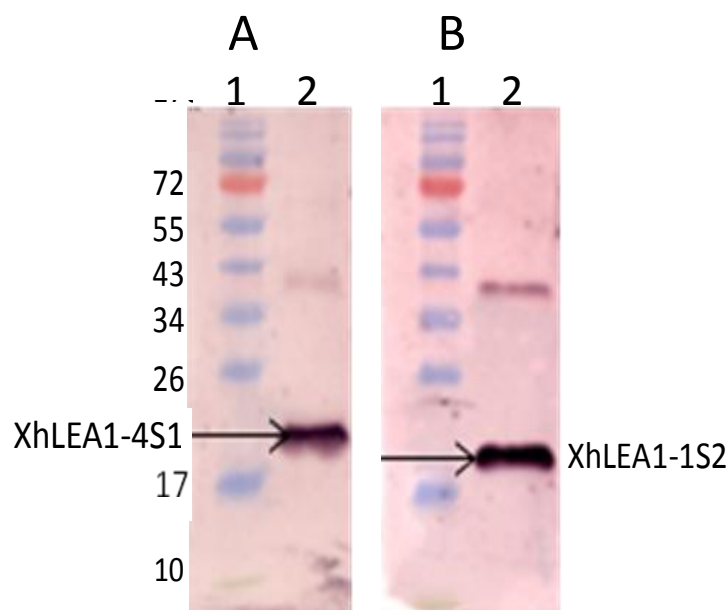


Figure 4.4 Confirmation of identities of *XhLEA1-4S1* and *XhLEA1-1S2* recombinant proteins by western blotting. Heat treated recombinant proteins expressed using pET-21a(+) vector carrying the respective LEA cDNAs were separated on 12 % SDS gel and identified by western blotting with antibodies generated against these proteins. Lane 1 in both A and B are for molecular weight marker (Prestained Protein Ladder, #SM0671), Lane 2 is for *XhLEA1-4S1* in A and *XhLEA1-1S2* in B.

The thermal stability test showed that both *XhLEA1-4S1* and *XhLEA1-1S2* recombinant proteins were heat stable, i.e. the proteins remained in solution. The bulk of the heating sensitive proteins were aggregated and removed by the centrifugation step. When the supernatant was analyzed using SDS PAGE, it showed the recombinant proteins plus few other protein bands some of which were of low concentration (Figure 4.5). This step resulted in partially purified *XhLEA1-4S1* and *XhLEA1-1S2* recombinant proteins and cleared the lysate improving the performance of the T7-Tag affinity purification column in the isolation of the specific XhLEA proteins. The heat stable and co-purified proteins were then removed by this last purification step resulting in pure recombinant proteins that were suitable for structural and functional studies (Figure 4.6).

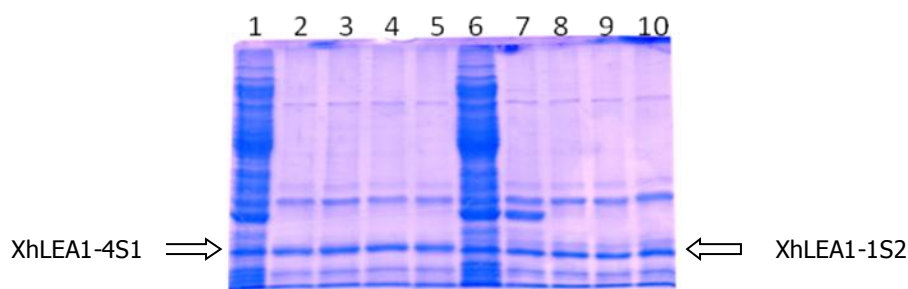


Figure 4.5 Thermal stability test of lysates containing *XhLEA1-4S1* and *XhLEA1-1S2* recombinant proteins. Lanes 1&6 represent untreated total protein lysate containing *XhLEA1-4S1* and *XhLEA1-1S2* recombinant proteins respectively; Lanes 2 to 5 and 7 to 10 are lysates containing *XhLEA1-4S1* and *XhLEA1-1S2* proteins respectively that are heated for 10 minutes at 70, 80, 90 and 100 °C respectively.

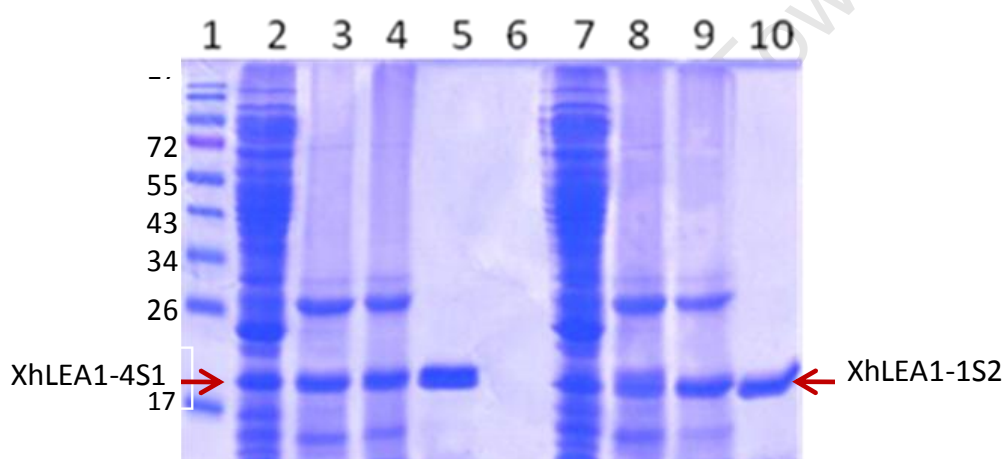


Figure 4.6 Purification of *XhLEA1-4S1* and *XhLEA1-1S2* recombinant proteins using T7 Tag Affinity purification kit. Lane 1= molecular weight marker (Prestained Protein Ladder, #SM0671), lanes 2 and 7 represent the total protein lysates containing *XhLEA1-4S1* and *XhLEA1-1S2* recombinant proteins respectively. Lanes 3 and 4 are of *XhLEA1-4S1*, and lanes 8 and 9 are of *XhLEA1-1S2* proteins partially purified by heating at 80 and 90 °C respectively. Lane 5 = purified *XhLEA1-4S1* and Lane 10 = purified *XhLEA1-1S2*. Lane 6 was not loaded and used as a separating lane.

4.3.2 Structural studies

PONDR algorithm predicted that both *XhLEA1-4S1* and *XhLEA1-1S2* were likely to be disordered proteins (Figure 4.7 A and B respectively). Proteins with a PONDR score of above 0.5 are classified as disordered and 149 residues of the total 152 of *XHLEA1-4S1* make up a continuous disordered region resulting in 98.03 % percent disorder with an average disorder score of 0.8079 (Figure 4.7 A). Similarly 120 residues of the total 136 residues of *XhLEA1-1S2* make up the disordered region resulting in 87.50 % disorder with

an average disorder prediction score of 0.7399. However the disorder in the latter LEA was not continuous and the disordered regions showed different disorder strength (0.7526 - 0.8292). The predicted ordered region in *XhLEA1-1S2* is 16 amino acids long located centrally stretching from 40 to 55 amino acid residues (Figure 4.7 B).

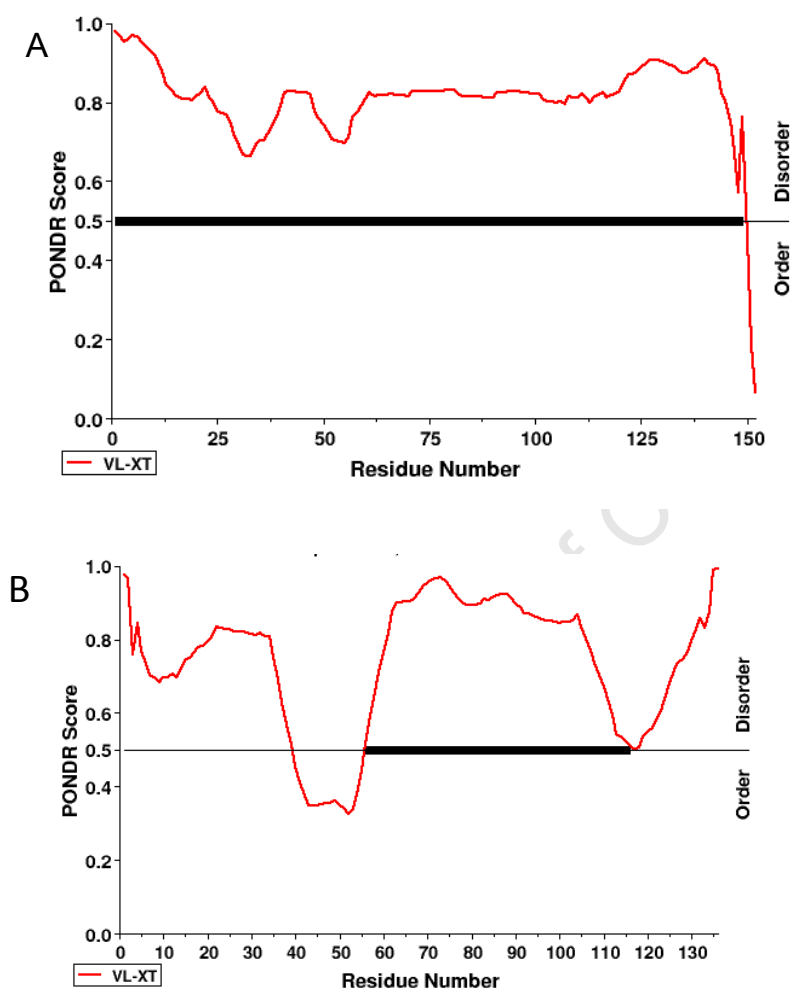


Figure 4.7 Prediction of structural disorder of *XhLEA1-4S1* (A) and *XhLEA1-1S2* (B) proteins using PONDR. PONDR scores of residues are indicated on the vertical axis against residue number on the horizontal axis. Residues making up a continuous disordered region are indicated by horizontal bold line.

Similarly, a Uversky plot of mean net charge against mean scaled hydropathy (Uversky 2002) indicates that *XhLEA1-4S1* and *XhLEA1-1S2* are intrinsically disordered proteins (Figure 4.8) with net charge and hydropathy scores of 0.0197 and 0.3398 for *XhLEA1-4S1*,

and 0.0147 and 0.3983 for *XhLEA1-1S2* respectively. These values are located close to the bottom hydropathy line (Figure 4.8, A and B) indicating that the proteins are weak in charge but highly hydrophilic. The grand average hydropathicity scores were found to be -1.441 and -0.915 for *XhLEA1-4S1* and *XhLEA1-1S2* respectively, the former being more hydrophilic.

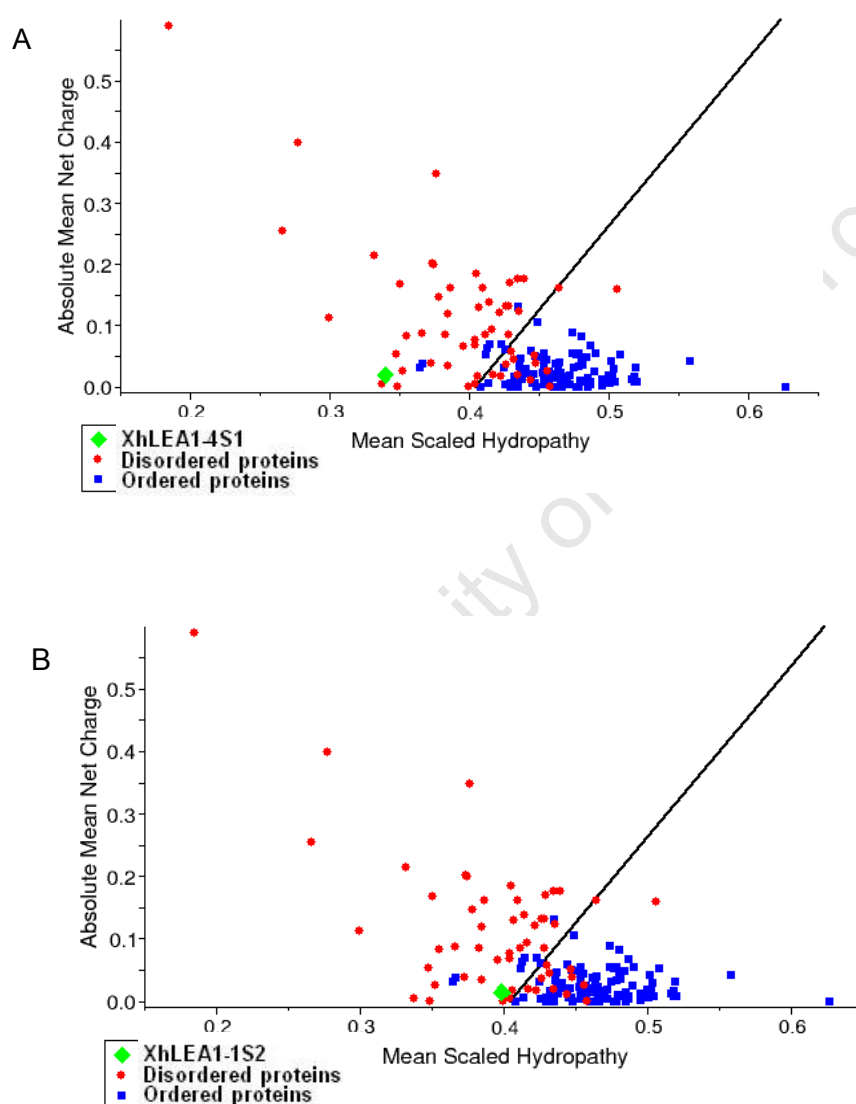


Figure 4.8 A Uversky plots of *XhLEA1-4S1* (A) and *XhLEA1-1S2* (B) amino acid sequences. *XhLEA1-4S1* and *XhLEA1-1S2* (green diamonds) are located with disordered proteins (red circles), whereas ordered set of proteins (blue squares) are found on the right side of the slant line dividing ordered and disordered proteins based on mean net charge and hydropathy scores.

The CD analysis of the affinity purified *XhLEA1-4S1* and *XhLEA1-1S2* proteins in PBS buffer showed a pattern similar to that of a random coil structure in the range of 185 – 260 nm wave length (Figure 4.9 A and B). CD runs were increased to smoothen the line patterns in all samples but did not bring much change. The jagged line pattern observed particularly with *XhLEA1-4S1* was related to the low concentration used in the study. The spectra of both XhLEA proteins showed a large minima at around 200 nm consistent with earlier reports on unstructured LEA proteins (Soulages *et al.*, 2003; Shih *et al.*, 2004 and 2010; Mouillon *et al.*, 2006) with a noticeable ellipticity occurring near zero around 210 and 220 nm wave length.

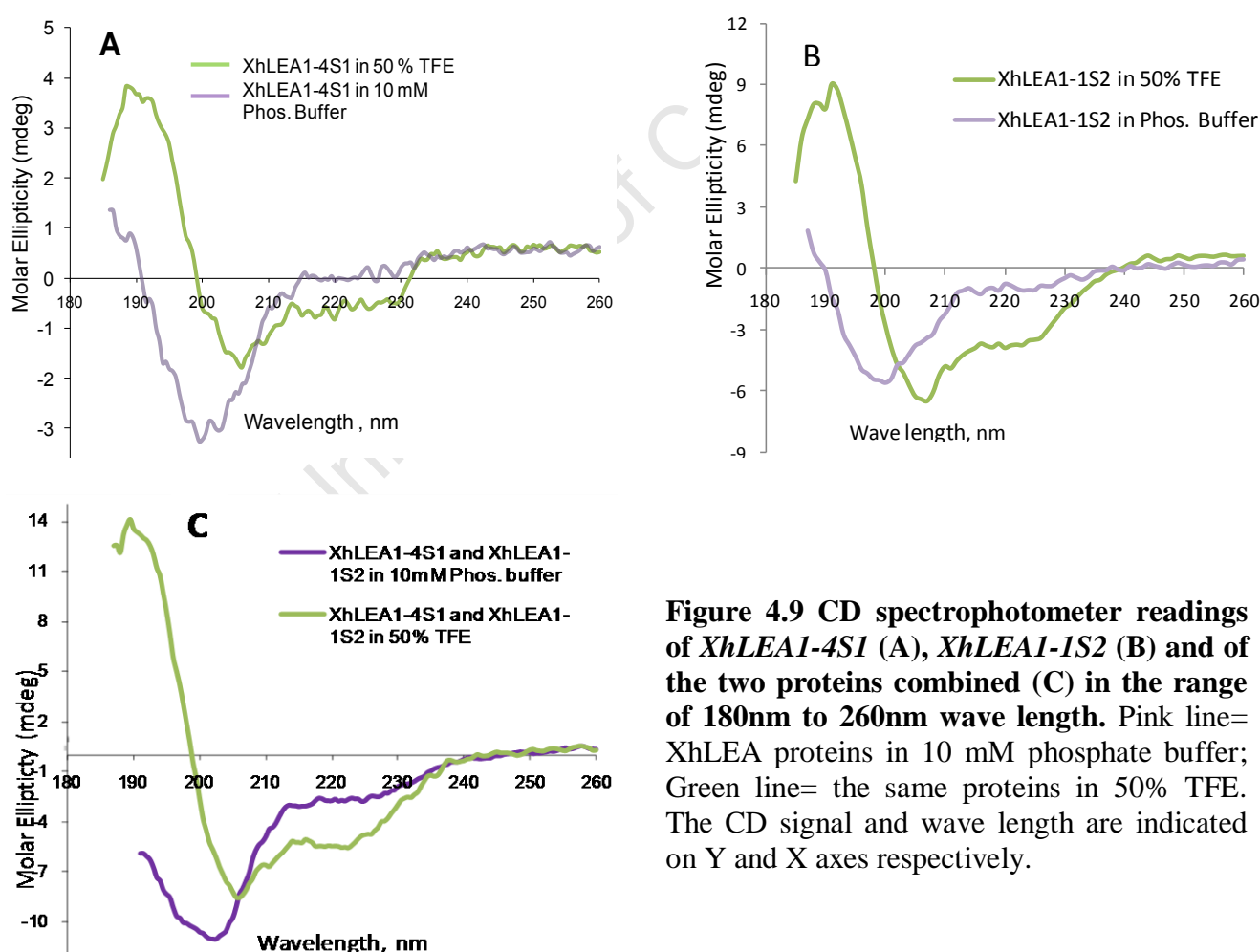


Figure 4.9 CD spectrophotometer readings of *XhLEA1-4S1* (A), *XhLEA1-1S2* (B) and of the two proteins combined (C) in the range of 180nm to 260nm wave length. Pink line= XhLEA proteins in 10 mM phosphate buffer; Green line= the same proteins in 50% TFE. The CD signal and wave length are indicated on Y and X axes respectively.

The objective of adding TFE was to investigate the intrinsic ability of these XhLEA proteins to form a secondary structure. Although the exact mechanism is not well understood, TFE is known to promote the formation of secondary structure. It has been proposed that this compound aggregates preferentially around proteins removing water molecules from the proximity and hence reinforcing hydrogen bonding between carbonyl and amide groups of the protein (Luo and Baldwin, 1997). On addition of 50 % TFE, a different spectra were observed for both *XhLEA1-4S1* and *XhLEA1-1S2* proteins (Figure 4.9 A and B). The large minima of the negative band shifted to around 207 and 208nm wave length for *XhLEA1-4S1* and *XhLEA1-1S2* respectively. Another broad shouldered negative band was observed to form at around 221nm for *XhLEA1-1S2* and for the combined *XhLEA1-4S1* and *XhLEA1-1S2*. The latter band was not clearly visible on the *XhLEA1-4S1* spectra which could be due to the high noise. On the positive side of all the graphs, one strong band was observed at around 190 nm wave length. The effect of the 16 amino acid stretch in *XhLEA1-1S2* predicted to be ordered could not be confirmed by CD analysis before and after the addition of TFE.

Deconvolution of the CD data obtained before and after the addition of TFE also showed that the α -helical content of *XhLEA1-4S1* and *XhLEA1-1S2* proteins increased by 45 % and 55 % on addition of TFE, respectively. There was also an apparent change in the shape of the spectral band when the two proteins were combined in the presence or absence of 50 % TFE (Figure 4.9 C), however, this did not show a significant difference in the overall percentage composition of the different secondary structures. The observed difference in the shape of the spectral band was likely to be related to the increased concentration of the protein in the solution. The CD results obtained were similar to the published data of unstructured LEA proteins (Shih *et al.*, 2004 and 2010a).

4.3.3 Anti aggregation role of *XhLEA1-4S1* and *XhLEA1-S2*

Aggregation and enzyme activity assays were conducted to investigate if *XhLEA1-4S1* and *XhLEA1-1S2* proteins have a stabilization role on the desiccation sensitive CS enzyme on drying. In a pilot experiment, the XhLEA proteins on their own did not show aggregation after two cycles of vacuum drying and rehydration. However, when the CS was dried and rehydrated in the same way in the absence of any protectant, it aggregated (Table 4.2). OD readings of dried CS increased as a function of concentration. The results of the pilot experiment showed that the XhLEA proteins did not contribute to the aggregation readings of CS in subsequent experiments. Furthermore, drying of CS by vacuum or Dura-Stop™ microprocessor-controlled tray drier did not show any difference in aggregation readings; hence the data reported in this study was obtained by vacuum drying.

Table 4.2 *XhLEA1-4S1* and *XhLEA1-1S2* aggregation test. A two-cycle vacuum drying and rehydration pilot experiment, each for two hours at room temperature, showed that *XhLEA1-4S1* and *XhLEA1-1S2* proteins did not aggregate on drying, but CS was aggregated. Untreated control samples were kept at 4 °C and the aggregation of two different concentration of CS, i.e. CS1 and CS2 in the absence of protectants were tested.

Sample	Amount dried, mg	OD@340nm after 2 drying cycles
Untreated <i>XhLEA1-4S1</i>	0.12	0.0176
Dried <i>XhLEA1-4S1</i>	0.12	0.0202
Untreated <i>XhLEA1-1S2</i>	0.12	0.0128
Dried <i>XhLEA1-1S2</i>	0.12	0.0178
Untreated CS	0.12	0.0016
Dried CS 1	0.12	0.2755
Dried CS 2	0.24	0.4006

CS enzyme was chosen to investigate the protein stabilization role of *XhLEA1-4S1* and *XhLEA1-1S2* as this particular protein/enzyme is known to aggregate on drying. OD readings of different CS aggregation experiments performed over a wide range of molar ratios showed that these XhLEA proteins are not effective in preventing CS from drying-induced aggregation. The result of a range of CS:LEA molar ratios tested are shown in

Figure 4.10 A, B and C. It was established that although some reduction in CS aggregation was observed at molar ratios of CS:XhLEA of 1:1 for *XhLEA1-4S1* and 1:5 for *XhLEA1-1S2*, this was not comparable to the protection provided by the Nematode LEA, *AavLEA1* (Figure 4.10 C).

Triplicate aggregation experiments, at the molar ratios of 1:1 for *XhLEA1-4S1* and 1:5 for *XhLEA1-1S2* that showed an apparent reduction in aggregation, were performed to see if the observed reduction in aggregation (Figure 4.10, A, B and C) was statistically significant and reproducible. In agreement with earlier reports (Goyal *et al.*, 2005; Chakrabortee *et al.* (2007), *AavLEA1* prevented CS aggregation to the levels of the untreated CS control at a molar ratio of LEA:CS of 1:5 ($p < 0.001$, Figure 4.10, D). Although, some reduction in CS aggregation was observed in the presence of the XhLEAs, this was not significantly different compared to the CS dried control and to provisional results obtained using BSA as a protectant instead of LEA (results not shown). BSA was previously reported to have no protective effect on CS aggregation (Goyal *et al.*, 2005).

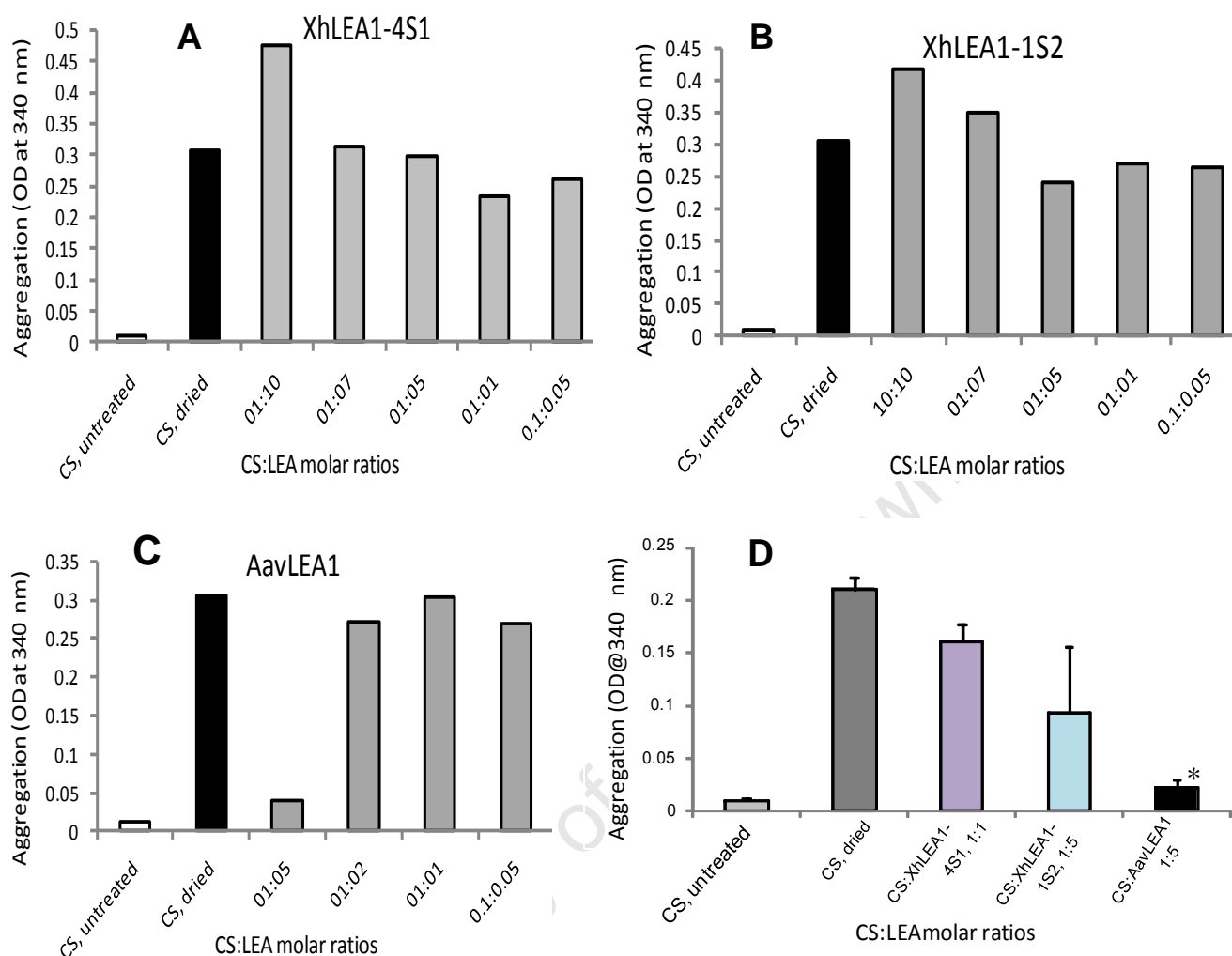


Figure 4.10. CS aggregation in the presence or absence of *XhLEA1-4S1* (A), *XhLEA1-1S2* (B) or *AavLEA1* (C). Molar ratios that showed apparent reduction in CS aggregation were repeated in triplicate (D). 0.12 mg of CS was dried and re-hydrated twice in the presence or absence of LEA protectants at the molar ratios indicated. Aggregation (light scattering) was measured as reported previously (Goyal *et al.*, 2005). One drying cycle corresponds to a drying of the mixtures followed by immediate rehydration in water to the original volume (200 μ l). Statistically significant results ($p < 0.001$) are indicated by ‘*’. Error bars (in D) indicate standard deviation

4.3.4 Enzyme stabilization role of *XhLEA1-4S1* and *XhLEA1-S2*

Enzyme activity results showed that CS activity was maintained to about 75 to 80 % of untreated control when *XhLEA1-4S1* was used at 1:1 and 1:5 molar ratio respectively, a significant protection when compared to the dried CS control (CS without LEA protein,

$p < 0.001$). Similar activity results were reported before when CS was dried in the presence of BSA, although BSA did not prevent CS aggregation (Goyal *et al.*, 2005). However, combination of CS:*AavLEA1* at a molar ratio of 1:5 has shown more enzyme stabilization compared to *XhLEA1-4S1*. On the other hand, activity results of CS dried in the presence of *XhLEA1-1S2* at a 1:1 and 1:5 CS:LEA molar ratio did not show significant difference when compared to dried CS control (Figure 4.11).

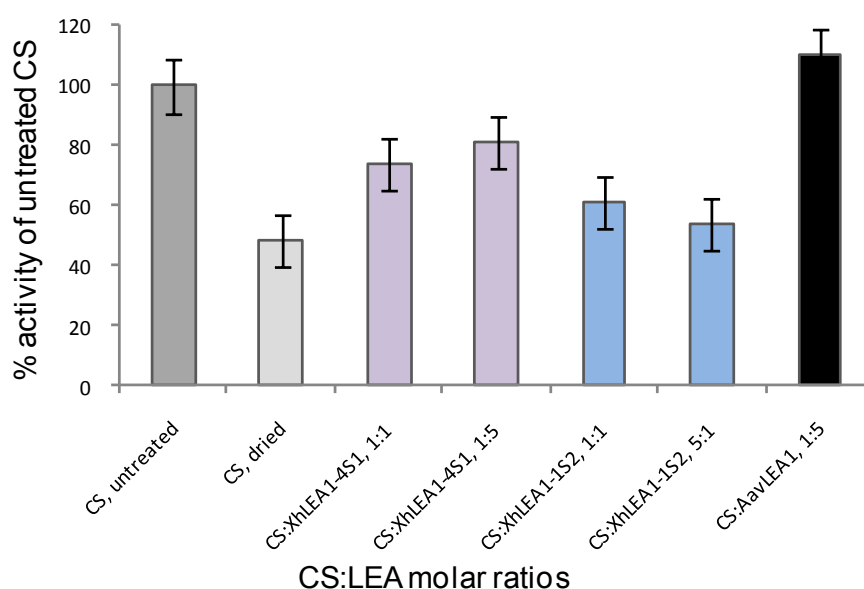


Figure 4.11 Effect of *XhLEA* proteins on the activity of 0.12mg desiccated CS enzyme. Activity assay was performed using a standard method on 2 μ l (of the 200 μ l total volume) of dried and re-hydrated CS sample in the presence or absence of *XhLEA1-4S1* (pink bars) or *XhLEA1-1S2* (blue bars) or *AavLEA1* (black bar) as described by Goyla *et al.* (2005). Untreated (deep gray bar) and dried (light gray) CS only controls were included. Enzyme activity results after two drying cycles are expressed as percentage of untreated CS control activity. Assays were repeated three times and results of CS:*XhLEA1-4S1* and CS:*AavLEA1* are statistically significant ($p < 0.001$). Error bars represent standard error.

4.4 Discussion

The heating step introduced into the purification procedure was a modification into the T7-Tag Affinity purification protocol. This was intended to minimize the concentration of unnecessary proteins and maximize the relative abundance of the recombinant proteins. In general, LEA proteins are heat stable and this heating step can be used to facilitate purification process from tissue samples or bacterial lysates. By removing all heat-sensitive proteins, this step improves the performance of subsequent affinity based purification columns. By increasing the relative abundance of the heat stable proteins in the solution, it also increases the surface area for the affinity interaction to take place. This step can be incorporated into the T7-Tag affinity purification protocol provided that the protein of interest is heat stable. Even heating at 50 °C can help in removing heat sensitive proteins and simplify further purification.

XhLEA1-IS2 migrated shorter distance on 12 % SDS gel than would be expected of its molecular weight (+T7-Tag = 15.2 kDa). A similar anomaly has been reported before for purified 13.1 kDa tomato ASR1 protein which migrated in SDS-PAGE as an 18 kDa protein (Kalifa *et al.*, 2004, Rom *et al.*, 2006). This anomaly was related to the high content of charged residues in ASR1. However, *XhLEA1-IS2* protein has low charge; hence the anomaly observed with *XhLEA1-IS2* could not be related to charge. Other factors reported to affect the migration of proteins on SDS include incomplete reduction, differences in SDS binding, degradation of the protein sample and inappropriate polyacrylamide concentration. Incomplete reduction and degradation are unlikely causes for the observed difference in migration. This is due to the fact that there are no Cys residues in these proteins to be reduced and degradation would not lead to an increased molecular weight on a 12 % SDS-PAGE. Therefore, difference in SDS binding might have resulted in the migration difference observed. In this study, the identity of the band appeared on the SDS gel was confirmed to

be of *XhLEA1-1S2* using specific antibodies generated against this protein, hence, the possible factors described were not investigated.

PONDR was successfully used to predict protein disorder in many previous studies (Obradovic *et al.*, 2005; Peng *et al.*, 2006; Fuxreiter *et al.*, 2007) and is reported to be the best-known tool for the prediction of intrinsic disorder in given proteins sequences (Oxford Protein Production Facility, 2007). PONDR-VL-XT predictor was selected from the available options as it integrates three neural networks trained with various settings (<http://pondr.com/pondr>). In agreement with many reports on other Group 1 LEA proteins, PONDR results showed that *XhLEA1-4S1* and *XhLEA1-1S2* proteins are unstructured, the former being entirely disordered and the latter showing about 90 % disorder.

PONDR structure prediction results were supported by CD spectroscopy results where the latter showed the presence of more random coil structure in both *XhLEA1-4S1* and *XhLEA1-1S2* proteins compared to α -helical or β -sheet content. However, the effect of the 16 amino acid long, ordered region shown by PONDR for *XhLEA1-1S2* was not clearly seen in CD results. Furthermore, the spectral bands for both *XhLEA1-4S1* and *XhLEA1-1S2* were not a perfect match to the typical curve of random coiled structure especially beyond 210 nm wave length. It has been reported that the CD data of LEA proteins in physiological buffers show some residual secondary structure which results in a spectra slightly different from that of an ideal pure random coil (Mouillon *et al.*, 2006). Therefore, the difference observed might indicate the presence of limited secondary structure content which might include the 16 amino acid long peptide in *XhLEA1-1S*.

Addition of TFE to *XhLEA1-4S1* and *XhLEA1-1S2* protein solutions was to reduce water availability and create a condition similar to what would be expected of a dehydrating

cytoplasm. After the addition of 50 % TFE, CD data produced a different curve similar to that of α -helical protein that might indicate the ability of these proteins to form α -helical structure. This feature is attributed to many other LEA proteins (Mouillon *et al.*, 2006 and 2008;; Gilles *et al.*, 2007; Tolleter *et al.*, 2007 and 2010; Kovacs, *et al.*, 2008; Soulages *et al.*, 2002; Shih *et al.*, 2010a; Shih *et al.*, 2010c).

The secondary structure of a given protein can also be determined by deconvoluting the data obtained from the CD machine. However, deconvolution data of intrinsically unstructured proteins such as LEAs on its own might not be reliable and can only be used to support results obtained by an alternative method. This is due to the fact that these softwares are developed using data bases of commonly available globular proteins and may not give an accurate picture of unstructured proteins. However, one can compare experimental CD spectra with the characteristic CD spectra of the three main protein structures (α -helical, β -sheet and random coil) and make reasonable conclusion (Professor Horst Klump, personal communication).

The results obtained showed that *XhLEA1-4S1* and *XhLEA1S2* proteins have intrinsic ability to form a secondary structure under limiting water activity. However, only few experimental reports suggest that such formation of secondary structure is related with function. Gilles *et al.* (2007), using enzyme protection assay, demonstrated that loss of α -helical domain of a wheat Em protein (Group 1 LEA) reduced its ability to protect LDH enzyme from drying-induced damage. On the other hand, LEA proteins are reported to resist forming secondary structure which caused some investigators to speculate that being an unstructured might be a requirement for their role in desiccation tolerance (Moulin *et al.*, 2008). This is supported by the fact that unstructured proteins are more flexible and can move easily in the cell milieu than proteins with secondary structure (Wright and Dyson,

19990). Chakrabortee and colleagues (2010) also found out that a nematode DNA binding anhydriin remained unstructured when bound to DNA. These findings show that more work needs to be done to establish LEA protein function.

In related studies, addition of osmolytes and crowding agents to LEAs did not contribute to change in structure when the mixture was analyzed with CD (Moulin *et al.*, 2008; Shih *et al.*, 2010a). In this study we investigated if the CD data of the combined *XhLEA1-4S1* and *XhLEA1-1S2* proteins was different from the scan data of each separately. There was no noteworthy difference between the two. However, this doesn't exclude the possibility that these proteins might play additive protective role against desiccation. Due to the scope of this project, such additive role was not investigated.

Increased aggregation when *XhLEA1-4S1* and *XhLEA1-1S2* proteins were dried alone and re-hydrated was not expected due to the high solubility and hydrophilic nature of the proteins. However, aggregation experiments have shown that both *XhLEA1-4S1* and *XhLEA1-1S2* proteins were not effective in preventing CS aggregation on drying. The effects observed by these XhLEA proteins on CS aggregation were not significantly different than when BSA was used instead. On the other hand, in agreement with previous reports (Goyal *et al.*, 2005), the aggregation assays clearly showed the anti-aggregation role of *AavLEA1* protein on dried CS in the same experiments. From the results obtained, it was concluded that *XhLEA1-4S1* and *XhLEA1-1S2* do not provide sufficient protection to other proteins against drying-induced aggregation.

However, *XhLEA1-4S1* and *XhLEA1-1S2* proteins are differentially expressed in response to dehydration in *X. humilis* (Chapter 2) and were found to be associated with plasma membrane associated with cell wall in dry leaves and roots (Chapter 3). Therefore, these

proteins might have other protective function against desiccation induced damage to the cell wall-plasma membrane-cytoplasm interface.

Enzyme activity of CS dried in the presence of *XhLEA1-4S1* was maintained, but not when it was dried in the presence of *XhLEA1-1S2*. The difference observed between these two proteins might be related to the difference in the number of 20-mer motif. This motif is repeated 4 times in *XhLEA1-4S1* and once in *XhLEA1-1S2*. If this motif is responsible for water binding property of Group 1 LEA proteins (Bray 1993; Close, 1996; Cuming 1999) then *XhLEA1-4S1* would be expected to provide more hydration buffering during drying or better facilitate hydration of molecules during rehydration. It has been shown previously that CS enzyme dried in the presence of BSA also maintained activity (Goyal *et al.*, 2005). These results indicate that aggregation of CS enzyme does not necessarily translate to loss of function and that aggregated CS enzyme might still have some enzymatically active sites for a residual enzyme activity, and the difference observed between the two XhLEA proteins might be related to the number of the 20-mer motif .

This study has demonstrated that *XhLEA1-4S1* and *XhLEA1-1S2* are typical Group 1 LEA proteins without a secondary structure in a neutral solution. However, they have shown intrinsic ability to form α -helical structure under decreased water activity, which could be an indication of what might happen in dehydrating plant tissues. Due to the extreme hydrophilic nature of LEA proteins, it is difficult to provide direct evidence to the implied concept. The current study also showed that the primary function of *XhLEA1-4S1* and *XhLEA1-1S2* proteins under desiccation stress, as evidenced by the *in vitro* aggregation assays, is unlikely to be protein stabilization. However, their expression in response to water loss (Chapter 2) is evidence that they have a role in desiccation tolerance. As *XhLEA1-4S1* and *XhLEA1-1S2* proteins were found associated with membrane structures (Chapter 3), their function is likely to be related to membrane stabilization and not metabolic processes.

Chapter 5

General conclusions

The association of LEA proteins with desiccation tolerance in resurrection plants and anhydrobiotic animals has been established over the last two decades. However, their precise role has not yet been defined. Although different functions have been proposed for many LEA proteins, experimental evidence is still limited. In the current study, expression, structure and location of *XhLEA1-4S1* and *XhLEA1-1S2*, Group 1 LEA proteins in leaves, roots and seeds of the resurrection plant *X. humilis* have been described. The results obtained indicate that *XhLEA1-4S1* and *XhLEA1-1S2* proteins are typical unstructured Group 1 LEA proteins and are expressed in response to desiccation stress in leaves and roots; and as part of the maturation drying process in seeds of *X. humilis*. The proteins were localized in the plasma membrane-cell wall compartment in dry leaves and dry roots, and in seeds they are found associated with protein storage vacuoles (PSVs).

One of the many proposed functions of LEA proteins is stabilization of other proteins against drying-induced aggregation or denaturation. From the results obtained, *XhLEA1-4S1* and *XhLEA1-1S2* proteins did not show effective enzyme/protein stabilization role. Another proposed function of LEA proteins is membrane stabilization. Although, no assay was performed to investigate the latter, the findings that *XhLEA1-4S1* and *XhLEA1-1S2* proteins are expressed in response to desiccation and that they were localized in the plasma membrane-cell wall structures of dry vegetative tissues indicate that these proteins may have a role in membrane stabilization. The absence of these proteins in the cytoplasm where enzymatic reactions are more prevalent, on the other hand, indicate that *XhLEA1-4S1* and *XhLEA1-1S2* proteins are not involved in stabilizing metabolic processes.

The proposed membrane stabilization function may be achieved through the formation of hydrogen bonding network or facilitation of, along with other sugars, glass formation

(verification) that could prevent loss of structure and membrane adhesion. Wolfe and Bryant (1999) described verification as occurring, due to the presence of solutes, between membranes to reduce stress and strains on membranes under reduced hydration due to freezing. Similar finding was made by Wolfe et al., (2002) where freezing point was depressed by the presence of hydrophilic solutes under conditions of low hydration. The fact that both drying and freezing reduce hydration, mechanisms of tissue protection under these stresses might be similar and the localization of *XhLEA1-4S1* and *XhLEA1-1S2* in the membrane area could be related to this function. The next logical step in the study of these proteins with respect to their proposed membrane protection function would be testing this function by membrane or liposome stability assay.

Furthermore, it is also possible that *XhLEA1-4S1* and *XhLEA1-1S2* proteins, due to their hydrophilic nature, provide a hydration buffering effect for other macromolecules during drying which was evidenced by the effect of *XhLEA1-4S1* on CS enzyme activity on drying. For desiccation-tolerant plants this hydration buffering effect could be important as it would allow slow drying which is required for the establishment of desiccation tolerance mechanisms.

It is worthwhile mentioning that the production and purification of antibodies, using the facilities provided, took considerable time. Having these antibodies made by commercial suppliers and spending more time on other aspects of the research would be advisable for future studies. The polyclonal antibodies used for determination of expression conditions and immunolocalization studies were raised against the GST-XhLEA proteins. It would be useful to test the antigenicity of the XhLEA proteins without the GST tag so that protein-specific antibodies could be used instead.

Further evidence of the role of *XhLEA1-4S1* and *XhLEA1-1S2* in desiccation tolerance could

be obtained by expressing these LEA proteins in bacteria or yeast. Furthermore, other studies on these proteins could be undertaken to find out if they are desiccation-specific by investigating their expression under other stresses such as freezing, osmotic or heat stresses. This study is the first systematic approach undertaken to understand the expression conditions, localization and function of *XhLEA1-4S1* and *XhLEA1-1S2* proteins in desiccation tolerance in the resurrection plant *X. humilis*. The results obtained will benefit future LEA protein characterization efforts and will contribute towards understanding the role of LEA proteins in desiccation tolerance.

University of Cape Town

References

1. **Alban C., Job D., Douce R. (2000).** Biotin metabolism in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 51:17–47.
2. **Ali-Benali M.A., Alary R., Joudrier P., AND Gautier M. F. (2005).** Comparative expression of five *Lea* genes during wheat seed development and in response to abiotic stresses by real-time quantitative RT-PCR. *Biochim. Biophys. Acta.* 1730: 56–65.
3. **Almoguera C., Jordano J. (1992).** Developmental and environmental concurrent expression of sun-flower seed stored low molecular weight heat shock protein and *Lea* mRNAs. *Plant Molecular Biology.* 19: 781-792.
4. **Alsheikh M.K., Heyen B.J., Randall S.K. (2003).** Ion binding properties of the dehydrin ERD14 are dependent upon phosphorylation. *J. Bio.l Chem.* 278:40882–40889.
5. **Babu R.C., Zhang J.S., Blum A., Ho T., Wu R., Nguyen H.T. (2004).** *HVA1*, a LEA gene from barley confers dehydration tolerance in transgenic rice (*Oryza sativa* L.) via cell membrane protection. *Plant Sci.* 166:855–862.
6. **Bahrndorff S., Tunnacliffe A., Wise M. J., McGee B., Holmstrup M., Loeschcke V. (2009).** Bioinformatics and protein expression analyses implicate LEA proteins in the drought response of Collembola. *J. Insect Physiol.* 55(3): 210-217.
7. **Baker J.C., Steele C., Dure L. III (1988).** Sequence and characterization of six *Lea* proteins and their genes from cotton. *Plant Mol. Biol.* 11:277—291.
8. **Bartels D., Salamini F. (2001).** Desiccation tolerance in the resurrection plant *Craterostigma plantagineum*. A contribution to the study of drought tolerance at the molecular level. *Plant Physiol* 127:1346–1353.
9. **Bartels D., Schneider K., Terstappen G., Piatkowski D., Salamini F. (1990).** .Molecular cloning of abscisic acid modulated genes which are induced during desiccation of the resurrection plant *Craterostigma plantagineum*. *Planta* 181:27–34.

10. **Battaglia M., Olvera-Carrillo Y., Garciarrubio A., Campos F., Covarrubias A.A. (2008).** The enigmatic LEA proteins and other hydrophilins. *Plant Physiol.* 148: 6– 24.
11. **Battista J.R., Park M.J., McLemore A.E. (2001).** Inactivation of two homologues of proteins presumed to be involved in the desiccation tolerance of plants sensitises *Deinococcus radiodurans* R1 to desiccation. *Cryobiology.* 43:133.
12. **Bies, N., Aspart, L., Carles, C., Gallois, P. and Delseny, M. (1998).** Accumulation and degradation of Em proteins in *Arabidopsis thaliana*: evidence for post-transcriptional controls. *J. Exp. Bot.* 49: 1925–1933.
13. **Bies-Etheve N., Gaubier-Comella P., Debures A., Lasserre E., Jobet E., Raynal M., Cooke R., Delseny M. (2008).** Inventory, evolution and expression profiling diversity of the LEA (late embryogenesis abundant) protein gene family in *Arabidopsis thaliana*. *Plant Mol. Biol.* 67, 107–124.
14. **Bishop R.F., Cipriani E., Lund J.S., Barnes G.L., Hosking C.S. (1984).** Estimation of rotavirus immunoglobulin G antibodies in human serum samples by enzyme-linked immunosorbent assay: expression of results as units derived from a standard curve. *J. Clin. Microbiol.* 19:447.
15. **Bockel C., Salamini F., Bartels D. (1998).** Isolation and characterization of genes expressed during early events of the dehydration process in the resurrection plant *Craterostigma plantagineum*. *J. Plant. Physiol.* 152:158–166.
16. **Borovskii G.B., Stupnikova I.V., Antipina A.I., Vladimirova S.V. and Voinikov V.K. (2002).** Accumulation of dehydrin-like proteins in the mitochondria of cereals in response to cold, freezing, drought and ABA treatment. *BMC Plant Biol.* 2: 1–7
17. **Boudet J., Buitink J., Hoekstra F.A., Rogniaux H., Larré C., Satour P., Leprince O. (2006).** Comparative Analysis of the Heat Stable Proteome of Radicles of *Medicago truncatula* Seeds during Germination Identifies Late Embryogenesis Abundant Proteins Associated with Desiccation Tolerance. *Plant Physiol.* 140:1418-1436.

18. **Bradford M.M. (1976).** Rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72: 248–254.
19. **Bray E.A. (1991).** Regulation of gene expression by endogenous ABA during drought stress. In: *Abscissic Acid: Physiology and Biochemistry*, (W.J. Davies, and H.G. Jones, Eds) Bios Scientific Publishers, Oxford, UK, 81-98.
20. **Bray E.A. (1993).** Molecular Responses to Water Deficit. *Plant Physiol.* 103:1035-1040.
21. **Bray E.A. (1994).** Alterations in Gene Expression in Response to Water Deficit. In: *Stress-Induced Gene Expression in Plants*. (A.S. Basra, ed.). Harwood Academic, 1-23.
22. **Bray E.A. (1997).** Plant responses to water deficit, *Trends Plant Sci.* 2:48–54.
23. **Brini F., Hanin M., Lumbreras V., Amara I., Khoudi H., Hassairi A., Pagès M., Masmoudi K. (2007b).** Overexpression of wheat dehydrin DHN-5 enhances tolerance to salt and osmotic stress in *Arabidopsis thaliana*. *Plant Cell Rep.* 26: 2017–2026.
24. **Brini F., Hanin M., Lumbreras V., Irar S., Pagès M., Masmoudi K. (2007a).** Functional characterization of DHN-5, a dehydrin showing a differential phosphorylation pattern in two Tunisian durum wheat (*Triticum durum* Desf.) varieties with marked differences in salt and drought tolerance. *Plant Sci.* 172:20–28.
25. **Browne J., Tunnacliffe A. Burnell A. (2002).** Plant desiccation gene found in a nematode. *Nature.* 416:38.
26. **Browne JA, Dolan KM, Tyson T, Goyal K, Tunnacliffe A, Burnell AM. (2004).** Dehydration-specific induction of hydrophilic protein genes in the anhydrobiotic nematode *Aphelenchus avenae*. *Eukaryot. Cell* 3:966–75.
27. **Campbell S.A., Close T.J. (1997).** Dehydrins: genes, proteins, and associations with phenotypic traits. *New Phytol.* 137 61–74.

28. **Campos F., Zamudio F. Covarrubias A.A. (2006).** Two different late embryogenesis abundant proteins from *Arabidopsis thaliana* contain specific domains that inhibit *Escherichia coli* growth. *Biochem. Biophys. Res. Commun.* 342(2):406-413.
29. **Cartwright T. (2001).** *Protein Purification Applications*, (ed. Simon Roe), Oxford University Press, AEA Tech, Oxford UK.
30. **Chakrabortee S., Boschetti C., Walton L.J., Sarkar S., Rubinsztein D.C., Tunnacliffe A. (2007).** Hydrophilic protein associated with desiccation tolerance exhibits broad protein stabilization function. *Proc. Natl. Acad. Sci. USA* 104:18073–18078.
31. **Chakrabortee S., Meersman F., Schierle G.S.K., Bertoncini C.W. McGee B., Kaminski C.F., Tunnacliffe A. (2010).** Catalytic and chaperone-like functions in an intrinsically disordered protein associated with desiccation tolerance. *PNAS*. 107(37):16084-16089.
32. **Chaves M.M., Maroco J.P. Pereira J.S., (2003).** Understanding plant responses to drought: from genes to the whole plant, *Funct. Plant Biol.* 30:239–264.
33. **Chen Z.Q., Liu Q., Zhu Y.S., Li Y.X., (2002).** Assessment on predicting methods for membrane-protein transmembrane regions, *Biochem. Biophys. Acta.(Sin)*. 34 (3):285–290.
34. **Cheng Z.Q., Targolli J., Huang X.Q., Wu R. (2002).** Wheat LEA genes PMA80 and PMA 1959, enhance dehydration tolerance of transgenic rice (*Oryza sativa* L.). *Mol. Breed.* 10:71-82.
35. **Classen D.C., Morningstar J.M., Shanley, J.D. (1987).** Detection of antibody to murine cytomegalovirus by enzyme-linked immunosorbent and indirect immunofluorescence assays. *J. Clin. Microbiol.* 25:600.
36. **Close T.J. (1996).** Dehydrins: Emergence of a biochemical role of a family of plant dehydration proteins. *Physiol. Plant.* 97:795–803.
37. **Close T.J. (1997).** Dehydrins: a commonality in the response of plants to dehydration and low temperature. *Physiol Plant.* 100:291–296.
38. **Close T.J., Fenton R.D., Moonan F. (1993).** A view of the plant dehydrins using antibodies specific to the carboxy terminal peptide. *Plant Mol. Biol.* 23:279–286.

39. **Close T.J., Lammers P. J. (1993).** An osmotic stress protein of cyanobacteria is immunologically related to plant dehydrins. *Plant Physiol.* (Bethesda) 101:773-779.
40. **Cohen A., Plant A.L, Moses M.S., Bray E.A. (1991).** Organ-specific and environmentally regulated expression of two abscisic acid-induced genes of tomato: nucleotide sequence and analysis of the corresponding cDNAs. *Plant. Physiol.* 97:1367–1374.
41. **Collett H., Shen A., Gardner M., Farrant J.M., Denby K.J., Illing N. (2004).** Towards transcript profiling of desiccation tolerance in *Xerophyta humilis*: Contraction of a normalized 11k *X. humilis* cDNA set and microarray expression analysis of 424 cDNAs in response to dehydration. *Physiol. Plant.* 122:39-53.
42. **Colmenero-Flores J.M., Campos F., Garcarrubio A., Covarrubias A.A. (1997).** Characterization of *Phaseolus vulgaris* cDNA clones responsive to water deficit: identification of a novel late embryogenesis abundant-like protein. *Plant Mol. Biol.* 35:393–405.
43. **Colmenero-Flores J.M., Mxoreno L.P., Smith C.E., Covarrubias A.A. (1999).** *Pvlea-18*, a member of a new late-embryogenesis-abundant protein family that accumulates during water stress and in the growing regions of well-irrigated bean seedlings. *Plant Physiol.* 120:93–104.
44. **Cuming A.C. (1999).** LEA proteins. In PR Shewry, R Casey, eds, Seed Proteins. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 753–780.
45. **Cuming A.C., Cho S.H., Kamisugi Y., Graham H., Quatrano R.S. (2007).** Microarray analysis of transcriptional responses to abscisic acid and osmotic, salt, and drought stress in the moss, *Physcomitrella patens*. *New Phytol.* 176:275–287.
46. **Cuming A.C., Lane B.G.. (1979).** Protein synthesis in imbibing wheat embryos. *Eur. J. Biochem.* 99:217--224.
47. **Curry J., Walker-Simmons M.K. (1993).** Unusual sequence of Group 3 LEA (II) mRNA inducible by dehydration stress in wheat. *Plant Mol. Biol.* 21:907-12.

48. **Dace H., Sherwin H.W., Illing N., Farrant J.M. (1998).** Use of metabolic inhibitors to elucidate mechanisms of recovery from desiccation stress in the resurrection plant *Xerophyta humilis*. *Plant Growth Regul.* 24:171-177.
49. **Dalal M., Tayal D., Chinnusamy V., Bansal K.C. (2009).** Abiotic stress and ABA-inducible Group 4 LEA from *Brassica napus* plays a key role in salt and drought tolerance. *J. Biotechnol.* 139:137–145.
50. **D'Amico F., Skarmoutsou E. (2008).** Clustering and colocalization in transmission immunoelectron microscopy: a brief review. *Micron* 39:1-6.
51. **Davidson S.W., Jonas A., Clayton D.F., George J.M. (1998).** Stabilization of alpha-Synuclein Secondary Structure upon Binding to Synthetic Membranes. *J. Biol. Chem.* 273:9443-9449.
52. **Delseny M., Bies-Etheve N., Carles C., Hull G., Vicient C., Raynal M., Grellet F., Aspart L. (2001).** Late Embryogenesis Abundant (LEA) protein gene regulation during *Arabidopsis* seed maturation. *J. Plant Physiol.* 158:419–427.
53. **Dure L. (1993).** A repeating 11-mer amino acid motif and plant desiccation. *Plant J.* 3:363–369.
54. **Dure L. (2001).** Occurrence of a repeating 11-mer amino acid sequence motif in diverse organisms. *Protein Pept. Lett.* 8:115–122.
55. **Dure L., Crouch M., Harada J., Ho T.-H.D., Mundy J., Quatrano R., Thomas T., Sung Z.R. (1989).** Common Amino Acid Sequence Domains among the LEA Proteins of Higher Plants. *Plant Mol. Biol.* 12:475-486.
56. **Dure L., Greenway S.C., Galau G.A. (1981).** Developmental biochemistry of cottonseed embryogenesis and germination. XIV. Changing mRNA populations as shown by *in vitro* and protein synthesis. *J. Biochem.* 20:4162–4168.
57. **Espelund M., Saeboe-Larssen S., Hughes D.W., Galau G.A., Larsen F., Jakobsen K.S. (1992).** Late embryogenesis-abundant genes encoding proteins with different numbers of hydrophilic repeats are regulated differentially by abscisic acid and osmotic stress. *Plant J.* 2:241–252.

58. **Farrant J.M. (2000).** A comparison of patterns of desiccation tolerance among three angiosperm resurrection plant species. *Plant Ecol.* 151: 29-39.
59. **Farrant J.M., Berjak P., Pammenter N.W. (1992).** Proteins in development and germination of a desiccation-sensitive (recalcitrant) seed species. *J. Plant Growth Regul.* 11: 257-265.
60. **Finkelstein R., Gampala S., Rock C. (2002).** Absciscic acid signaling in seeds and seedlings. *Plant Cell.* 14:S15–S45.
61. **Fisher K.M. (2008).** Bayesian reconstruction of ancestral expression of the LEA gene families reveals propagule-derived desiccation tolerance in resurrection plants. *Am. J. Botany*, 95(4): 506 - 515.
62. **Frangioni J.V., Neel B.G. (1993).** Solubilization and purification of enzymatically active glutathione *S*-transferase (pGEX) fusion proteins. *Anal. Biochem.* 210:179–187.
63. **Frey A., Di Canzio J., Zurakowski D. (1998).** A statistically defined endpoint titer determination method for immunoassays. *J. Immunol. Methods.* 221(1-2):35-41.
64. **Futcher B., Latter G.I., Monardo P., McLaughlin C.S., Garrels J.I. (1999).** A sampling of the yeast proteome. *Mol. Cell Biol.* 19:7357–7368.
65. **Fuxreiter M., Tompa P., Simon I. (2007).** Local structural disorder imparts plasticity on linear motifs. *Bioinformatics* 23(8): 950–6.
66. **Galau G.A., Dure L. (1981).** Developmental biochemistry of cottonseed embryogenesis and germination: changing messenger ribonucleic acid populations as shown by reciprocal heterologous complementary deoxyribonucleic acid-messenger ribonucleic acid hybridization. *Biochem.* 20:4169–4178.
67. **Galau G.A., Hughes D.W., Dure. L. (1986).** Absciscic acid induction of cloned cotton late embryogenesis-abundant (Leu) mRNAs. *Plant Mol. Biol.* 7:155-170.
68. **Galau G.A., Wang H.Y., Hughes D.W. (1992).** Cotton *Lea4* (D19) and *LeaA2* (D132) Group 1 *Lea* genes encoding water stress-related proteins containing a 20-amino acid motif. *Plant Physiol.* 99:783–788

69. **Galau G.A., Wang H.Y-C., Hughes D.W. (1993).** Cotton *Leu5* and *Lea4* encode atypical late embryogenesis-abundant proteins. *Plant Physiol.* 101:695-696.
70. **Garay-Arroyo, A., Colmenero-Flores J.M., Garcarrubio A., Covarrubias A.A. (2000).** Highly hydrophilic proteins in prokaryotes and eukaryotes are common during conditions of water deficit. *J. Biol. Chem.* 275:5668–5674.
71. **Gathercole D.V., Colling D.J., Skepper J.N., Takagishi Y., Levi A.J., Severs N.J. (2000).** Immunogold-labeled L-type calcium channels are clustered in the surface plasma membrane overlying junctional sarcoplasmic reticulum in guinea-pig myocytes—implications for excitation–contraction coupling in cardiac muscle. *J. Mol. Cell Cardiol.* 32, 1981–1994.
72. **Gaubier P., Raynal M., Hull G., Huestis G.M., Grellet F., Arenas C., Pages M., Delseny M., (1993).** Two different Em-like genes are expressed in *Arabidopsis thaliana* seeds during maturation, *Mol. Gen. Genet.* 238:409–418.
73. **Gilles G.J., Hines K.M., Manfre A.J., Marcotte W.R. (2007).** A predicted N-terminal helical domain of a Group 1 LEA protein is required for protection of enzyme activity from drying. *Plant Physiol. Biochem.* 45:389-399
74. **Godoy J.A., Lunar R., Torres-Schumann S., Moreno J., Rodrigo R.M., Pintor-Toro J.A. (1994).** Expression, tissue distribution and subcellular localization of dehydrin TAS14 in salt-stressed tomato plants. *Plant Mol. Biol.* 26:1921–1934.
75. **Goldgur Y., Rom S., Ghirlando R., Shkolnik D., Shadrin N., Konrad Z., Bar-Zvi D, (2007).** Desiccation and zinc binding induce transition of tomato abscisic acid stress ripening 1, a water stress- and salt stress-regulated plant-specific protein, from unfolded to folded state. *Plant Physiol.* 143:617–628.
76. **Goyal K., Tisi L., Basran A., Browne J., Burnell A., Zurdo J., Tunnacliffe A. (2003).** Transition from natively unfolded to folded state induced by desiccation in an anhydrobiotic nematode protein. *J. Biol. Chem.* 278:12977–12984.
77. **Goyal K., Walton L. J., Tunnacliffe A. (2005).** LEA proteins prevent protein aggregation due to water stress. *Biochem. J.* 388,151 -157.
78. **Grelet J., Benamar A., Teyssier E., Avelange-Macherel M-H., Grunwald D., Macherel D. (2005).** Identification in pea seed mitochondria of a late-

embryogenesis abundant protein able to protect enzymes from drying. *Plant Physiol.* 137:157–167.

79. **Griffiths G., Hoppeler H. (1986).** Quantitation in immunocytochemistry: correlation of immunogold labeling to absolute number of membrane antigens. *J. Histochem. Cytochem.* 34:1389–1398.
80. **Griffiths G., Lucocq J.M., Mayhew T.M. (2001).** Electron microscopy applications for quantitative cellular microbiology: *Technoreview, Cellular Microbiology.* 3(10):659-668.
81. **Gygi S.P., Rochon Y., Franza B.R., Aebersold R. (1999).** Correlation between protein and mRNA abundance in yeast. *Mol Cell Biol.* 19:1720–1730.
82. **Han B., Hughes D.W., Galau G.A., Bewley J.D., Kermode A.R. (1997).** Changes in late-embryogenesis-abundant (LEA) messenger RNAs and dehydrins during maturation and premature drying of *Ricinus communis* L. seeds. *Planta.* 201:27-35.
83. **Han B., Kermode A.R. (1996).** Dehydrin-like proteins in castor bean seeds and seedling are differentially produced in response to ABA and water-deficit-related stresses, *J. Exp. Bot.* 42:933–939.
84. **Hancock D.C., O'Reilly N.J. (2004).** Synthetic Peptides as Antigens for Antibody Production. In: *Methods in Molecular Biology, Immunochemical Protocols, Human Press* 295:13-25.
85. **Hand S.C., Jones D., Menze M.A., Witt T.L. (2007).** Life without water: expression of plant LEA genes by an anhydrobiotic arthropod. *J. Exp. Zool.* 307A:62–66.
86. **Hand S.C., Menze M.A., Toner M., Boswell L., Moore D. (2010).** LEA Proteins During Water Stress: Not Just for Plants Anymore. *Annu. Rev. Physiol.* 73:115-134.
87. **Hanly W.C., Artwohl J.E., Bennett B.T. (1995).** Review of polyclonal antibody production procedures in mammals and poultry. *ILAR News* 37:93.

88. **Hara M., Terashima S., Fukaya T., Kuboi T. (2003).** Enhancement of cold tolerance and inhibition of lipid peroxidation by citrus dehydrin in transgenic tobacco. *Planta* 217:290–298.
89. **Harada J.J., DeLisle A.J., Baden C.S., Crouch M.L. (1989).** Unusual sequence of an abscisic acid-inducible mRNA which accumulates late in *Brassica napus* seed development. *Plant Mol. Biol.* 12:395 – 401.
90. **He J.X., Fu J.R. (1996).** The research progresses in Lea proteins of seeds. *Plant Physiol. Commun.* 32 (4):241–246.
91. **Heazlewood J.L., Tonti-Filippini J., Verboom R.R., Millar A.H. (2005).** Combining experimental and predicted datasets for determination of subcellular location of proteins in *Arabidopsis*. *Plant Physiol.* 29:598–609.
92. **Heddererson N., Balsamo R., Cooper K., Farrant J.M. (2009).** Leaf tensile properties of resurrection plants differ among species in their response to drying. *South Afri. J. Bot.* 75: 8-16.
93. **Hellwege E.M., Dietz K.J., Hartung W. (1996).** Absciscic acid causes changes in gene expression involved in the induction of the landform of the liverwort *Riccia fluitans* L. *Planta* 198: 423–432.
94. **Heyen B.J., Alsheikh M.K., Smith E.A., Torvik C.F., Seals D.F., Randall S.K. (2002).** The calcium-binding activity of a vacuole-associated, dehydrin-like protein is regulated by phosphorylation. *Plant Physiol.* 130: 675–687.
95. **Hoekstra F.A., Golovina E.A., Buitink J. (2001).** Mechanisms of plant desiccation tolerance. *Trends Plant Sci.* 6:431–438.
96. **Houde M., Daniel C., Lachapelle M., Allard F., Laliberté S., Sarhan F. (1995).** Immunolocalization of freezing-tolerance-associated proteins in the cytoplasm and nucleoplasm of wheat crown tissues. *Plant J.* 8:583–593.
97. **Houde M., Danyluk J., Laliberte J.F., Rassart E., Dhindsa R.S. and Sarhan F. (1992).** Cloning, characterization and expression of a cDNA encoding a 50 kilodalton protein specifically induced by cold acclimation in wheat. *Plant Physiol.* 99:1381-1387.

98. **Hsing Y.C., Chen Z.Y., Shih M.D., Hsieh J.S., Chow T.Y. (1995).** Unusual sequences of Group 3 LEA mRNA inducible by maturation or drying in soybean seeds. *Plant Mol. Biol.* 29:863–868.
99. **Illing N., Denby K.J., Collett H., Shen A., Farrant J.M. (2005).** The signature of seeds in resurrection plants: A molecular and physiological comparison of desiccation tolerance in seeds and vegetative tissues. *Integr. Comp. Biol.*, 45:771–787.
100. **Imai R., Zhang L., Ohta A., Bray E.A., Takagi M. (1996).** A lea-class gene of tomato confers salt and freezing tolerance when expressed in *Saccharomyces cerevisiae*. *Gene*. 170:243–248.
101. **Ingram J., Bartels D. (1996).** The molecular basis of dehydration tolerance in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 47:377–403.
102. **Iturriaga G., Schneider K., Salamini F., Bartels D. (1992).** Expression of desiccation-related proteins from the resurrection plant *Craterostigma plantagineum* in transgenic tobacco. *Plant Mol. Biol.* 20:555–558.
103. **Jensen A.B., Goday A., Figureueras M., Jessop A.C., Pagès M. (1998).** Phosphorylation mediates the nuclear targeting of the maize Rab17 protein. *Plant J.* 13:691–697.
104. **Jiang X.; Wang Y. (2004).** β -Elimination coupled with tandem mass spectrometry for the identification of in vivo and in vitro phosphorylation sites in maize dehydrin DHN1 protein. *Biochem.* 43:15567–15576.
105. **Kalifa Y, Pearlson E, Gilad A, Konrad Z, Scolnik PA, Bar-Zvi D (2004b)** Overexpression of the water and salt stress-regulated *Asr1* gene confers an increased salt tolerance. *Plant Cell Environ.* 27:1459–1468
106. **Kalifa Y., Gilad A., Konrad Z., Zaccai M., Scolnik P.A., Bar-Zvi D. (2004a).** The water- and salt-stress-regulated *Asr1* (abscisic acid stress ripening) gene encodes a zinc-dependent DNA-binding protein. *Biochem. J.* 381:373–378.
107. **Karlson D.T., Fujino T., Kimura S., Baba K., Itoh T., Ashworth E.N. (2003).** Novel plasmodesmata association of dehydrin-like proteins in cold-acclimated red-osier dogwood (*Cornus sericea*). *Tree Physiol.* 23: 759–767.

108. **Kiernan J. A. (2000).** Formaldehyde, formalin, paraformaldehyde and glutaraldehyde: What they are and what they do. *Microscopy Today*. 1: 8-12.
109. **Kikawada T., Nakahara Y., Kanamori Y., Iwata K., Watanabe M., McGee B., Tunnacliffe A., Okuda T. (2006).** Dehydration-induced expression of L proteins in an anhydrobiotic chironomid. *Biochem. Biophys. Res. Commun.* 348:56–61.
110. **Koag M.C., Wilkens S., Fenton R.D., Resnik J., Vo E., Close T.J. (2009).** The K-Segment of Maize DHN1 Mediates Binding to Anionic Phospholipid Vesicles and Concomitant Structural Changes. *Plant Physiol.* 150(3):1503-1514.
111. **Koag M-C, Fenton R.D., Wilkens S., Close T.J. (2003).** The binding of Maize DHN1 to Lipid Vesicles. Gain of Structure and Lipid Specificity. *Plant Physiol.* 131(1):309-316.
112. **Kovacs D., Agoston B., Tompa P. (2008).** Disordered plant LEA proteins as molecular chaperones. *Plant Signaling & Behavior*. 3(9):710-713.
113. **Lan Y., Cai D., Zheng Y.Z. (2005).** Expression in *Escherichia coli* of three different soybean late embryogenesis abundant (LEA) genes to investigate enhanced stress tolerance. *J. Int. Plant Biol.* 47:613–621.
114. **Leprince O., Buitink J. (2010).** Desiccation tolerance: From genomics to the field, *Plant Sci.*, 179:554–564.
115. **Li N.Y., Gao J.F., Wang P.H., (1998).** Protein characters induced by water stresses in wheat seedlings, *Acta Phytophysiol. Sin.* 24(1):65–71.
116. **Lin C., Guo W.W., Everson E., Thomashow M.F. (1990).** Cold acclimation in *Arabidopsis* and wheat: a response associated with expression of related genes encoding ‘boiling-stable’ polypeptides. *Plant Physiol.* 94: 1078–1083.
117. **Lin C., Thomashow M.F. (1992).** DNA sequence analysis of a complementary DNA for cold-regulated *Arabidopsis* gene *cor15* and characterization of the COR15 polypeptide. *Plant Physiol.* 99:519–525.
118. **Lisse T., Bartels D., Kalbitzer H.R., Jaenicke R. (1996).** The recombinant dehydrin-like desiccation stress protein from the resurrection plant *Craterostigma*

plantagineum displays no defined three-dimensional structure in its native state. *Biol. Chem.* 377:555–561.

119. **Litts J.C., Erdman M.B., Huang N., Karrer E.E., Noueiry A., Quatrano R.S., Rodriguez R.L. (1992).** Nucleotide sequence of the rice (*Oryza sativa*) *Em* protein gene (*Emp1*). *Plant Mol. Biol.* 19:335-337.
120. **Liu X., Wang Z., Wang L., Wu R., Phillips J., Deng X. (2009).** *LEA* 4 group genes from the resurrection plant *Boea hygrometrica* confer dehydration tolerance in transgenic tobacco. *Plant Sci.* 176:90-98.
121. **Longhi S., Receveur-Brechot V., Karlin D., Johansson K., Darbon H., Bhella D., Yeo R., Finet S. and Canard B. (2003).** The C-terminal domain of the measles virus nucleoprotein is intrinsically disordered and folds upon binding to the C-terminal moiety of the phosphoprotein. *J. Biol. Chem.* 278:18638–18648.
122. **Love J.J., Li X., Chung J., Dyson H.J., Wright P.E. (2004).** The LEF-1 high-mobility group domain undergoes a disorder-to-order transition upon formation of a complex with cognate DNA. *Biochem.* 43:8725-8734.
123. **Luo P., Baldwin R. L. (1997).** Mechanism of helix induction by trifluoroethanol: a framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water. *Biochemistry.* 36:8413–8421.
124. **Manfre A.J., LaHatte G.A., Climer C.R., Marcotte W.R. Jr. (2009).** Seed dehydration and the establishment of desiccation tolerance during seed maturation is altered in the *Arabidopsis thaliana* mutant *atem6-1*. *Plant Cell Physiol.* 50:243-253.
125. **Manfre A.J., Lanni L.M., Marcotte Jr., W.R. (2006).** The *Arabidopsis thaliana* Group 1 LEA Protein ATEM6 is Required for Normal Seed Development. *Plant Physiol.* 140:140-149.
126. **Marttila S., Santén K. (2007).** Practical aspects of immunomicroscopy on plant material. In: Mendéz-Vilas A, Díaz j (eds) Modern Research and Educational Topics in Microscopy. *Microscopy Series* 3(2):1015-1021.
127. **Marttila S., Tenhola T., Mikkonen A. (1996).** A barley (*Hordeum vulgare* L) LEA3 protein, HVA1, is abundant in protein storage vacuoles. *Planta.* 199: 602–611.

128. **Mayhew T. (2005).** How to Count your Gold: A Tutorial on TEM Immunogold Label Quantification, School of Biomedical Sciences, University of Nottingham, UK.
http://www.microscopy-analysis.com/files/jwiley_microscopy/2005_Mar_Mayhew.pdf
129. **Mayhew T. M., Griffiths G., Lucocq J.M. (2004).** Applications of an efficient method for comparing immunogold labelling patterns in the same sets of compartments in different groups of cells. *Histochem. Cell Biol.* 122:171-177.
130. **Mercado-Pimentel M.E., Jordan N.C., Aisemberg G.O. (2002).** Affinity purification of GST fusion proteins for immunohistochemical studies of gene expression, *Protein Expr. Purif.* 26:260–265.
131. **Mishler B.D., Churchill, S.P. (1985).** Transition to a land flora: phylogenetic relationships of the green algae and bryophytes. *Cladistics* 1:305-28.
132. **Moons A., De Keyser A., Van Montagu M. (1997).** A Group 3 LEA cDNA of rice, responsive to abscisic acid, but not to jasmonic acid, shows variety-specific differences in salt stress response. *Gene* 191:197–204.
133. **Morris, P.C., Kumar A., Bowles D.J., Cuming A.C. (1990).** Osmotic stress and abscisic acid induce expression of the wheat *Em* genes. *Eur. J. Biochem.* 190:625—630.
134. **Mouillon J.M., Eriksson S.K., Harryson P. (2008).** Mimicking the plant cell interior under water stress by macromolecular crowding: disordered dehydrin proteins are highly resistant to structural collapse. *Plant Physiol.* 148 1925–1937.
135. **Mouillon J.M., Gustafsson P., Harryson P. (2006).** Structural investigation of disordered stress proteins: comparison of full-length dehydrins with isolated peptides of their conserved segments. *Plant Physiol.* 141:638–650.
136. **Mundy J., Chua N-H. (1988).** Absciscic acid and water-stress induce the expression of a novel rice gene. *EMBO J.* 7:2279–2286.
137. **NDong C., Danyluk J., Wilson K.E., Pocock T., Huner N.P.A., Sarhan F. (2002).** Cold-regulated cereal chloroplast late embryogenesis abundant-like proteins. Molecular characterization and functional analyses. *Plant Physiol.* 129:1368–1381.

138. **Neuhaus J.M., Rogers J.C. (1998).** Sorting of proteins to vacuoles in plant cells. *Plant Mol. Biol.* 38:127–144.
139. **Ngubane A. (2008).** Cloning and characterization of LEA1-EM genes in the resurrection plant *X. humilis*. MSc. Thesis, University of Cape Town
140. **Nieslanik B.S., Atkins W.M. (2000).** The catalytic Tyr-9 of glutathione S-transferase A1-1 controls the dynamics of the C terminus. *J. Biol. Chem.* 275: 17447–**17451**.
141. **Nylander M., Svensson J., Palva E.T., Welin B.V. (2001).** Stress-induced accumulation and tissue-specific localization of dehydrins in *Arabidopsis thaliana*. *Plant Mol Biol.* 45: 263–279.
142. **Obradovic Z., Peng K., Vucetic S., Radivojac P., Dunker A.K. (2005).** Exploiting heterogeneous sequence properties improves prediction of protein disorder. *Proteins.* 61(Suppl 7):176-182.
143. **Olave-Concha N., Ruiz-Lara S., Muñoz X., Bravo L.A., Corcuera L.J. (2004).** Accumulation of dehydrin transcripts and proteins in response to abiotic stresses in *Deschampsia antarctica*. *Antarct Sci.* 16:175–184.
144. **Olvera-Carrillo Y., Campos F., Reyes J.L., Garcarrubio A., Covarrubias A.A. (2010).** Functional Analysis of the Group 4 Late Embryogenesis Abundant proteins reveals their relevance in the adaptive response during water deficit in *Arabidopsis thaliana*. *Plant Physiol.* 154:373-390
145. **Oxford Protein Production Facility (2007).** *PONDR Disorder Prediction Tool*, Bioinformatics Services, University of Oxford.
146. **Pelah D., Shoseyov O., Altman A. (1995).** Characterization of BspA, a major boiling stable, water-stress-responsive protein in aspen (*Populus tremula*). *Tree Physiol.* 15:673-678
147. **Peng K., Radivojac P., Vucetic S., Dunker A.K., Obradovic Z. (2006).** Length-dependent prediction of protein intrinsic disorder. *BMC Bioinformatics.* 7:208.
148. **Pow D.V., Wright L.L., Vaney D.I. (1995).** The immunocytochemical detection of amino-acid neurotransmitters in paraformaldehyde-fixed tissues. *J. Neurosci. Methods* 56:115–123.

149. **Puhakainen T., Hess M.W., Mäkelä P., Svensson J., Heino P., Palva E.T. (2004).** Overexpression of multiple dehydrin genes enhances tolerance to freezing stress in *Arabidopsis*. *Plant Mol.Biol.* 54(5):743–53.
150. **Rabbani M., Maruyama K., Abe H., Khan M., Katsura K., Ito Y., Yoshiwara K., Seki M., Shinozaki K., Yamaguchi-Shinozaki K.. (2003).** Monitoring Expression Profiles of Rice Genes under Cold, Drought, and High-Salinity Stresses and Absciscic Acid Application Using cDNA Microarray and RNA Gel-Blot Analyses 1 [w]. *Plant Physiol.* 133:1755–1767.
151. **Rajesh S., Manickam A. (2006).** Prediction of functions for two LEA proteins from mung bean. *Bioinformation* 1(4): 133-138.
152. **Raynal M., Depigny D., Cooke R., Delseny M. (1989).** Characterization of a radish nuclear gene expressed during late seed maturation. *Plant Physiol.* 91:829-836.
153. **Rensing S.A., Lang D., Zimmer A.D., Terry A., Salamov A., et al. (2008).** The Physcomitrella genome reveals evolutionary insights into the conquest of land by plants. *Science.* 319(5859):64–69.
154. **Reynolds E.S. (1963).** The use of lead citrate at high pH as an electron opaque stain for electron microscopy. *J. Cell Biol.* 17:208–212.
155. **Rinne P.L., Kaikuranta P.L., van der Plas L.H., van der Schoot C. (1999).** Dehydrins in cold-acclimated apices of birch (*Betula pubescens* Ehrh.): production, localization and potential role in rescuing enzyme function during dehydration. *Planta.* 209:377–388.
156. **Roberts J. K., DeSimone N. A., Lingle W. L., Dure L. (1993).** Cellular concentrations and uniformity of cell-type accumulation of two LEA proteins in cotton embryos. *Plant Cell.* 5:769–780.
157. **Rorat T., Szabala B.M., Grygorowicz W.J., Wojtowicz B., Yin Z., Rey P. (2006).** Expression of SK3-type dehydrin in transporting organs is associated with cold acclimation in *Solanum* species. *Planta.* 224: 205–221.
158. **Roth J., Bendayan M., Orci L. (1978).** Ultrastructural localization of intracellular antigens by the use of protein A-gold complex. *J. Histochem. Cytochem.* 26:1074–1081.

159. **Rüdiger A., Jorgensen P.L., Antranikian G. (1995).** Isolation and characterization of a heat-stable pullulanase from the hyperthermophilic archaeon *Pyrococcus woesei* after cloning and expression of its gene in *Escherichia coli*. *Appl. Environ. Microbiol.* 61:567–575.
160. **Russouw P.S., Farrant J.M., Brandt W.F., Maeder D., Lindsey G.G. (1995).** Isolation and characterisation of a heat soluble protein from pea (*Pisum sativum*) embryos. *Seed Sci. Res.* 5(1):137–144.
161. **Rybicki, E. (1979).** *The serology of the Bromoviruses*. M.Sc. Thesis, University of Cape Town, South Africa.
162. **Saavedra L., Svensson J., Carballo V., Izmendi D., Welin B., Vidal S. (2006).** A dehydrin gene in *Physcomitrella patens* is required for salt and osmotic stress tolerance. *The Plant J.* 45(2):237–49.
163. **Saito S., Hirai N., Matsumoto C., Ohigashi H., Ohta D., Sakata K., Mizutani M. (2004).** *Arabidopsis* CYP707As encode (+)-abscisic acid 8'-hydroxylase, a key enzyme in the oxidative catabolism of abscisic acid. *Plant Physiol.* 134:1439–1449.
164. **Samarah N. H., Mullen R. E., Cianzio S. R., and Scott P. (2006).** Dehydrin-Like Proteins in Soybean Seeds in Response to Drought Stress during Seed Filling *Crop Sci.*, 46: 2141–2150.
165. **Scarano A., Lezzi G. and Piattelli A. (2003).** Common Fixatives in Hard-Tissue Histology. *In: Hand book of histology methods for bone and cartilage*, Y.H. An and K.L. Martin (eds.) Human Press, 159-165.
166. **Sherwin H.W., Farrant J.M. (1996).** Differences in rehydration of three desiccation tolerant angiosperm species. *Annals of Botany* 78:703–710.
167. **Sherwin H.W., Farrant J.M. (1998).** Protection mechanisms against excess light in the resurrection plants *Craterostigma wilmsii* and *Xerophyta viscosa*. *Plant Growth Regulation.* 24:203-210.
168. **Shih M.D., Hoekstra , F.A. and Hsing , Y.I.C. (2008).** Late embryogenesis abundant proteins. *Adv. Bot. Res.* 48:211–255.

169. **Shih M.D., Hsieh T.Y., Lin T.P., Hsing Y.I.C., Hoekstra F.A. (2010a).**
Characterization of Two Soybean (*Glycine max* L.) LEA IV Proteins by Circular Dichroism and Fourier Transform Infrared Spectrometry *Plant Cell Physiol.* 51 (3):395-407.
170. **Shih M.D., Huang L.T., Wei F.J., Wu M.T., Hoekstra F.A., and Hsing Y.I.C. (2010c).** *OsLEA1a*, a New Em-Like Protein of Cereal Plants. *Plant Cell Physiol.*, 51(12):2132-2144.
171. **Shih M.D., Kao M.H., Lin S.C., Shu T.F., Hsieh K.L., Chung M.C., Tsou T.H., Hsing Y.I.C., Hsieh J.S. (2010b).** Tissue- and cellular localization of soybean (*Glycine max* L.) seed maturation protein transcripts. *Botanical Studies* 51:183-194.
172. **Shih M.D., Lin S.C., Hsieh J.S., Tsou C.H., Chow T.Y., Lin T.P., et al. (2004).** Gene cloning and characterization of a soybean (*Glycine max* L.) LEA protein, GmPM16. *Plant Mol. Biol.* 56:689-703.
173. **Shimizu T., Kanamori Y., Furuki T., Kikawada T., Okuda T., Takahashi T., Mihara H., Sakurai M. (2010).** Desiccation-Induced Structuralization and Glass Formation of Group 3 Late Embryogenesis Abundant Protein Model Peptides. *Biochem.* 49:1093–1104.
174. **Singh J., Whitwill S., Lacroix G., Douglas J., Dubuc E., Allard G., Keller W., Scherthaner J.P. (2009).** The use of Group 3 LEA proteins as fusion partners in facilitating recombinant expression of recalcitrant proteins in *E. coli*. *Protein Expression and Purification* 67:15–22.
175. **Small I., Peeters N., Legeai F., Lurin C. (2004).** Predotar: A tool for rapidly screening proteomes for N-terminal targeting sequences. *Proteomics* 4:1581–1590.
176. **Songyang Z., Shoelson S. E., McGlade J., Olivier P., Pawson T., Bustelo X. R., Barbacid M., Sabe H., Hanafusa H., Yi T. (1994).** Specific motifs recognized by the SH2 domains of Csk, 3BP2, fps/fes, GRB-2, HCP, SHC, Syk, and Vav. *Mol. Cell Biol.* 14(4): 2777–2785.

177. **Soulages J.L., Kim K., Arrese E.L., Walters C., Cushman J.C. (2003).**
Conformation of a Group 2 late embryogenesis abundant protein from soybean: evidence of poly (L-proline)-type II structure. *Plant Physiol.* 131:963–975.
178. **Soulages J.L., Kim K., Walters C., Cushman J.C. (2002).** Temperature-induced extended helix/random coil transitions in a Group 1 late embryogenesis-abundant protein from soybean. *Plant Physiol.* 128:822–832.
179. **Spurr A. R. (1969).** A low-viscosity epoxy resin embedding medium for electron microscopy. *J. ultrastruct. Res.* 26:31-43.
180. **Stacy R.A., Aalen R.B. (1998).** Identification of sequence homology between the internal hydrophilic repeated motifs of Group 1 late-embryogenesis-abundant proteins in plants and hydrophilic repeats of the general stress protein GsiB of *Bacillus subtilis*. *Planta* 206:476–478.
181. **Sun G.R., Peng Y.Z., Yang X.F., Zhang R. (2003).** The effect of drought stress upon protein component in leaves of *Betula platyphylla* seedlings, *Sci. Silvae Sin.* 39 (4):151–154.
182. **Swire-Clark G.A., Marcotte W.R. (1999).** The wheat LEA protein Em functions as an osmoprotective molecule in *Saccharomyces cerevisiae*. *Plant Mol. Biol.*, 39(1):117-128.
183. **Tao H., Liu W., Simmons B.N., Harris H.K., Cox T.C., Massiah M.A. (2010).**
Purifying natively folded proteins from inclusion bodies using sarkosyl, Triton X-100, and CHAPS. *Biotechniques.* 48(1):61-4.
184. **Taylor R.G., Walker D.C., McInnes R.R. (1993).** *E. coli* host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing. *Nucleic Acids Res.* 21:1677-1678.
185. **Thalhammer A., Hundertmark M., Popova A.V., Seckler R., Hinch D.K. (2010).**
Interaction of two intrinsically disordered plant stress proteins (COR15A and COR15B) with lipid membranes in the dry state. *Biochim Biophys Acta.* 1798(9):1812-1820.
186. **Thomashow M.F. (1999).** Plant cold acclimation: freezing tolerance genes and regulatory mechanisms, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 50:571–599.

187. **Tolleter D., Hinch D.K., Macherel D. (2010).** A mitochondrial late embryogenesis abundant protein stabilizes model membranes in the dry state. *Biochim. Biophys. Acta.* 1798 (10):1926-1933.
188. **Tolleter D., Jaquinod M., Mangavel C., Passirani C., Saulnier P., Manon S., Teyssier E., Payet N., Avelange-Macherel M.H. and Macherel D. (2007).** Structure and function of a mitochondrial late embryogenesis abundant protein are revealed by desiccation. *Plant Cell.* 19:1580–1589.
189. **Tompa P., Banki P., Bokor M., Kamasa P., Kovacs D., Lasanda G., Tompa K. (2006).** Protein-water and protein-buffer interactions in the aqueous solution of an intrinsically unstructured plant dehydrin: NMR intensity and DSC aspects. *Biophys J.* 91:2243–2249.
190. **Tompa P., Csermely P. (2004).** The role of structural disorder in the function of RNA and protein chaperones. *Faseb. J.* 18:1169-1175.
191. **Tompa P., Kovacs D. (2010).** Intrinsically disordered chaperones in plants and animals. *Biochem Cell Biol.* 88:167-174.
192. **Tompa P., Szasz C., Buday L. (2005).** Structural disorder throws new light on moonlighting. *Trends Biochem. Sci;* 30:484-9.
193. **Tunnacliffe A., Lapinski J., McGee B. (2005).** A putative LEA protein, but no trehalose, is present in anhydrobiotic bdelloid rotifers. *Hydrobiologia* 546:315–321.
194. **Tunnacliffe A., Wise M. J. (2007).** The continuing conundrum of the LEA proteins. *Naturwissenschaften.* 94:791-812.
195. **Uversky V.N. (2002).** What does it mean to be natively unfolded? *Eur. J. Biochem.* 269(1):2-12.
196. **Veeranagamallaiah G., Prasanthi J., Reddy K.E., Pandurangaiah M., Babu O.S., Chinta S. (2010).** Group 1 and 2 LEA protein expression correlates with a decrease in water stress induced protein aggregation in *horsegram* during germination and seedling growth. *J Plant Physiol.*
doi:10.1016/j.jplph.2010.09.007.

197. **Vicre M., Sherwin H.W., Driouich A., Jaffer M., Jauneau A., Farrant J.M. (1999).** Cell wall properties of hydrated and dry leaves of the resurrection plant *Craterostigma wilmsii*. *J. Plant Physiol.* 155:719-726.
198. **Wang G.T., Hu Z.H., Chen J.Q. (2003).** CBF transcriptional activators and their roles in enhancing stress tolerance, *Plant Physiol. Commun.* 39(4):402–410.
199. **Warner A.H., Miroshnychenko O., Kozarova A., Vacratsis P.O., MacRae T.H., Kim J., Clegg J.S. (2010).** Evidence for multiple Group 1 late embryogenesis abundant proteins in encysted embryos of *Artemia* and their organelles. *J. Biochem.* 148 (5): 581-592.
200. **Watson B.S., Asirvatham V.S., Wang L., Sumner L.W. (2003).** Mapping the proteome of Barrel Medic (*Medicago truncatula*). *Plant Physiol.* 131:1104–1123.
201. **Whitsitt M.S., Collins R.G., Mullet J.E. (1997).** Modulation of Dehydration Tolerance in Soybean Seedlings' Dehydrin Mat Is Induced by Dehydration but Not by Absciscic Acid. *Plant Physiol.* 114: 917-925.
202. **Williams B., Tsang A. (1991).** A maize gene expressed during embryogenesis is abscisic acid-inducible and highly conserved. *Plant Mol. Biol.* 16:916-923.
203. **Williamson J.D., and Quatrano R.S. (1988).** ABA-regulation of two classes of embryo-specific sequences in mature wheat embryos. *Plant Physiol.* 86:208-215.
204. **Wise M. J. (2003).** LEAping to conclusions: A computational reanalysis of late embryogenesis abundant proteins and their possible roles. *BMC Bioinformatics* 4: 52–70.
205. **Wise M.J. (2002).** The POPPs: clustering and searching using peptide probability profiles. *Bioinformatics* (Suppl 1) 18: S38–S45.
206. **Wise M.J., Tunnacliffe A. (2004).** POPP the question: what do LEA proteins do? *Trends Plant Sci* 9:13–17.
207. **Wolfe J., Bryant G., Koster K.L. (2002).** What is 'unfreezable water', how unfreezable is it and how much is there? *CryoLetters.* 23:157-166.
208. **Wolfe J., Bryant G. (1999).** Freezing, drying and/or vitrification of membrane-solute-water systems. *Cryobiology.* 39:103-129.

209. **Wolkers W.F., McCready S., Brandt W.F., Lindsey G.G., Hoekstra F.A. (2001).** Isolation and characterization of a D-7 LEA protein from pollen that stabilizes glasses *in vitro*. *Biochim. Biophys. Acta*. 1544:196–206.
210. **Wright P.E., Dyson H.J. (1999).** Intrinsically unstructured proteins: reassessing the protein structure–function paradigm, *J. Mol. Biol.* 293:321–331.
211. **Xu D., Duan X., Wang B., Hong B., Ho T-HD., Wu R. (1996).** Expression of a late embryogenesis abundant protein gene, *HVA1*, from barley confers tolerance to water deficit and salt stress in transgenic rice. *Plant Physiol.* 110:249–257.
212. **Yang C.Y., Chen Y.C., Jauh G.Y., Wang C.S. (2005).** A lily ASR protein involves abscisic acid signaling and confers drought and salt resistance in *Arabidopsis*. *Plant Physiol.* 139:836–846.
213. **Yang Q.H., Ye W.H., Song S.Q., Yin S.H. (2002).** Seed desiccation tolerance and its relationship to seed types and developmental stages, *Acta Bot. Boreal-Occident Sin.* 22(6):1518–1525.
214. **Yu J.N. (2003).** Cloning and characterization of tolerant stress genes from *Triticum aestivum* dissertation for Doctor's degree of Northwestern Sci-tech University of Agriculture and Forestry, Yangling, China.
215. **Yu J.N., Zhang J.S., Shan L., Chen S.Y. (2005).** Two new Group 3 LEA genes of wheat and their functional analysis in yeast. *J Integr. Plant Biol.* 47: 1372–1381.
216. **Zhang L., Ohta A., Takagi M., Imai R. (2000).** Expression of plant Group 2 and Group 3 lea genes in *Saccharomyces cerevisiae* revealed functional divergence among LEA proteins. *J Biochem.* 127:611–616.
217. **Zhao Y., Yun J., Shi F., Wang J., Yang Q., Chao Y. (2010).** Molecular cloning and characterization of a Group 3 LEA gene from *Agropyron mongolicum* Keng. *Afr. J. Biotechnol.* 9 (37):6040-6048.
218. **Zheng H., von Mollard G.F., Kovaleva V., Stevens T.H., Raikhel N.V. (1999).** The Plant Vesicle-associated SNARE AtVTI1a Likely Mediates Vesicle Transport from the Trans-Golgi Network to the Prevacuolar Compartment., *Mol. Biol. Cell.* 10:2251-2264.

Appendices

Appendix A: Full length nucleotide sequence of *XhLEA1-4S1* and *XhLEA1-1S2*

➤ *XhLEA1-4S1*

ATGGCTTCCCATCAAGAAAGGGCTGAGATGGACCGTAGGGCCAGGG
AAGGTGAGACTGTCGTACCCGGTGGTACCGGCGGAAAGAGCCTCGA
AGCTCAACAACACCTCGCCGAAGGTCGTAGCAAGGGAGGGCAGAC
GAGGAAGGAACAGTTGGGGACAGAGGGGTATCAGGAGATGGGCCA
CAAAGGTGGTGAGACGAGGAGGGAACAGCTGGGTTCAGAAGGGTA
CAGTGAAATCGGGCACAAAGGAGGGGAGGCGAGGAGAGAGCAACT
GGGTCTCTGAAGGGTACAGTGAAATCGGGCACAAAGGTGGTGAGAC
GAGGAGGGATCAGATTGGGTCGGAGGGATACCGTGAGATGGGTCG
CAAAGGTGGTCTGTCTGACCAAGGACGAGTCCGGGGGAGAACGTGCT
GCCCAGGAGGGCATAGAGATCGACGAGTCCAAGTACAGGACCAAT
GTT

➤ *XhLEA1-1S2*

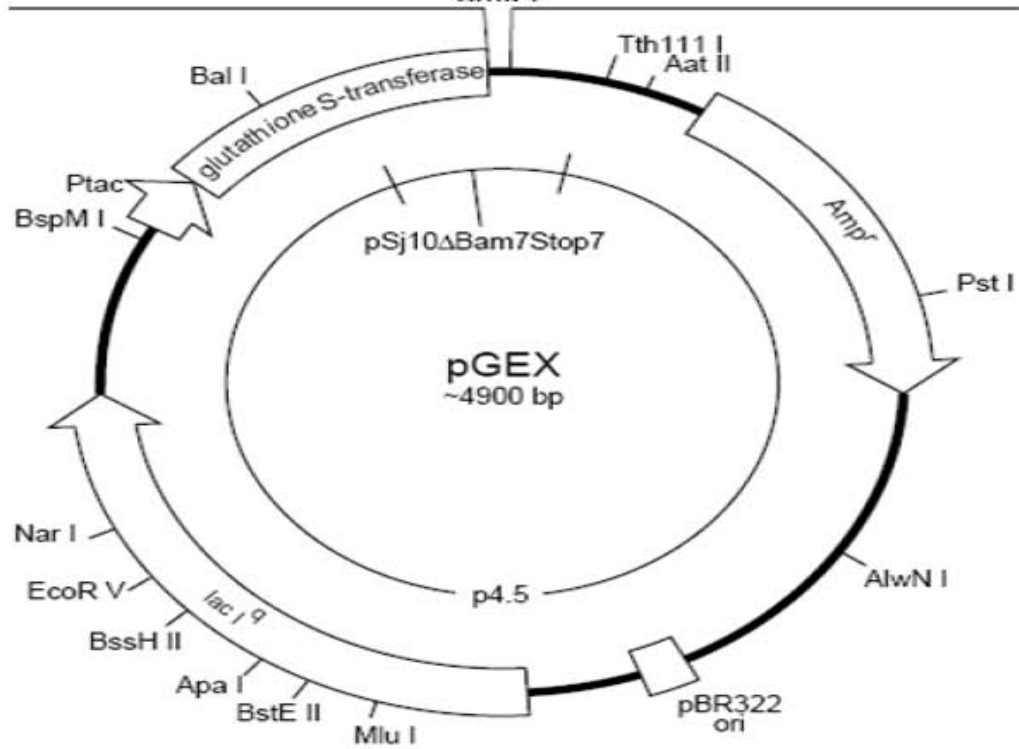
ATGGCTTCTGCACAGGAGAGGATTGAGCTCGACCGAAGGGCGAGG
GAGGGCGAGACTGTTGTCCCCGGCGGCACAGGCGGAAAGAGCCTC
GAAGCTCAAGAACACCTCGCTGACGGACGGAGCCGTGGAGGGCAG
ACGCGCAGGGATCAGCTGGGATCAGAAGGGTACAGTGAGCTTGGCC
GCATGGGTGGACAGAGCGGCGGCTCGTATTCCGATGAGACGGCCAC
CGGCGGGGTGTTGGGGCCCCGATCTCGGAGCAGGGGAGACGGTCGTA
CCTGGAGGCGCCGGCGGGAAGAGTGTTGAAGCTCAAGAGAATCTCG
CCGGAGGACGGAGACGCGGTGGGGAGACACGCATAGAGCAGCTTG
GTGCCGAAGGCTACAGTGCACTCGGCCGCACGGGGGGGAAGCAGC

Appendix B: pGEX-3X vector map

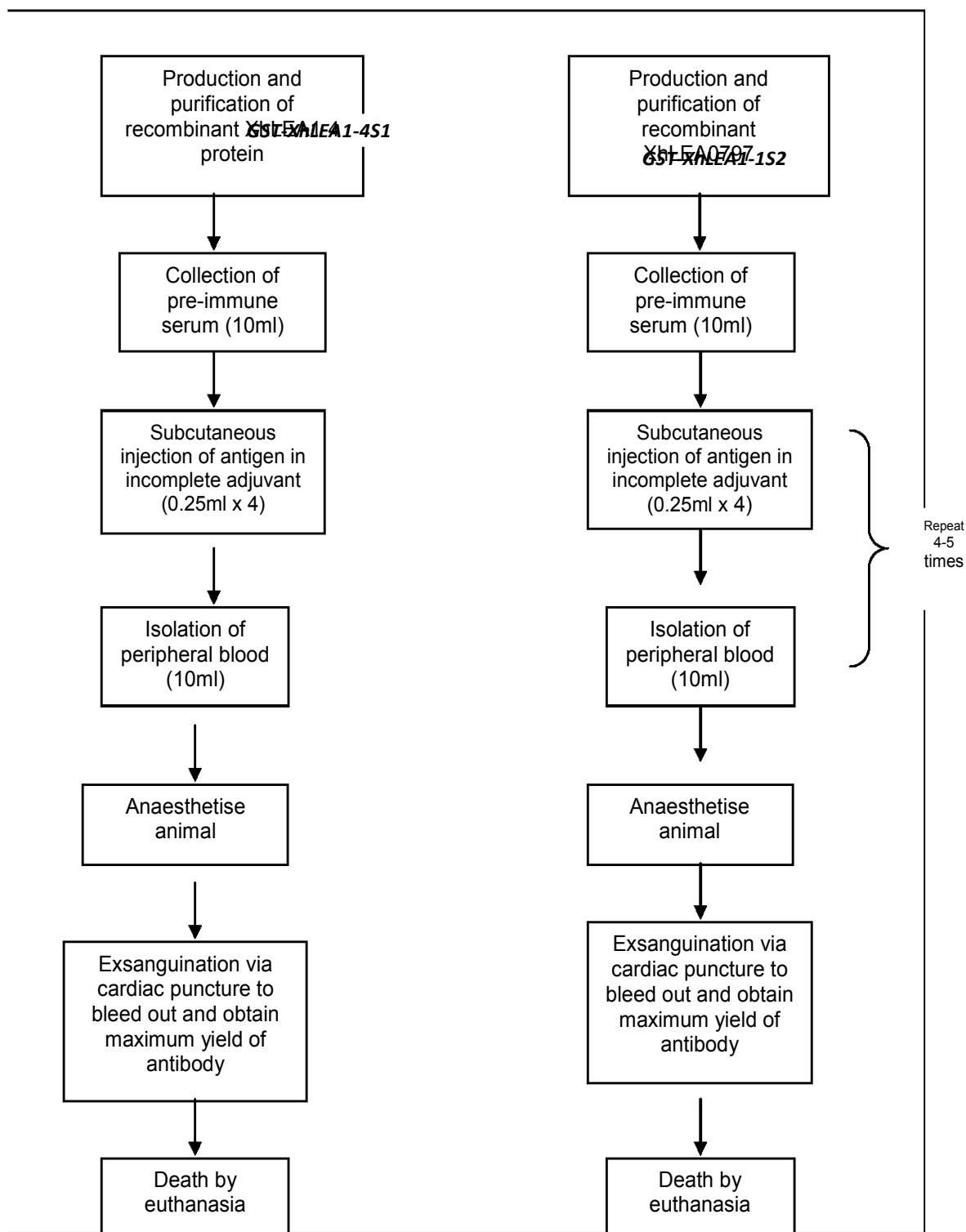
pGEX-3X (27-4803-01)

Factor Xa

Ile	Glu	Gly	Arg	Gly	Ile	Pro	Gly	Asn	Ser	Ser				
ATC	GAA	GGT	CGT	GGG	ATC	CCC	GGG	AAT	TCA	TCG	TGA	CTG	ACT	GAC
				BamH I		Sma I		EcoR I			Stop codons			



Appendix C: Rabbit immunization protocol



Appendix D: Alignment of Group 1 LEA proteins from Angiosperms and Moss. Conserved motifs that are repeated in the classical Group 1

LEAs (Tunnacliffe LEA motif_1) or which are repeated in the *X. humilis* XhLEA1-1S2 (PS00431) are indicated.

10 20 30 40 50 60 70 80 90 100
 LEA1 motif (PS00431)) -----GETVVPGGTGGKSLAQENLAEGRSR
 LEA1 motif_1 (Tunnacliffe) -----GGQTRREQLGEEGYSQMGRK
 LEA1 motif_1 (Tunnacliffe) -----GGQTRREQLGEEGYSQMGRK
 LEA1 motif_1 (Tunnacliffe) -----GGQTRREQLGE
 LEA1 motif_1 (Tunnacliffe) -----
 NM_115040_AT3G51810_A_thaliana MASKQLS--REELDEKAKQGETVVPGGTGGHSLEAQEHLAEGRSK---GGQTRKEQLGHEGYQEIGHKGGEARKEQLGHEGYQEMGHKGGEARKEQLGH
 XhLEA1-4S1 (B1) X humilis MASHQE---RAEMDRRAREGETVVPGGTGGKSLAQOHLAEGRSK---GGQTRKEQLGTEGYQEMGHKGGETRREQLGSEGYSEIGHKGGEARREQLGS
 Q05191_LEA_B19.4_Hordeum_vulga MASGQQE--RSELDRMAREGETVVPGGTGGKTLAQEHLAEGRSR---GGQTRKEQLGEEGYREMGHKGGETRKEQLGEEGYREMGHKGGETRKEQLGE
 BAD22769_LEA_[Bromus_inermis] MAFGQQE--RSELDRMAREGETVVPGGTGGKTLAQEHLAEGRSR---GGQTRKEQLGEEGYKEMGRKGGETRKEQLGEEGYKEMGRKGGETRKEQLGE
 XP_002439615_LEA_09g016830_[So MASGQDS--REELDRMAEEGQTVVPGGTGGKSLAQEHLAEGRSRSHGSGKGGQTRKEQLGHEGYVEMGKKGQTRSEQLGHEGYQEMSGKGGQTRREQLGH
 Q02400_LEA_B19.3_Hordeum_vulga MASGQQE--RSELDRMAREGETVVPGGTGGKTLAQEHLAEGRSR---GGQTRKDAQLGEEGYREMGHKGGETRKEQLGEEGYREMGHKGGETRKEQMGE
 ACG33377_LEA-1_[Zea_mays] MASGQES--REELARMAEEGQTVVPGGTGGKTLAQEHLAEGRSRSH---GGQTRSEQLGHEGYSEMGSKGGQTRKEQLGHEGYSEMGRK-----
 NP_001105349_LEA-5_[Zea_mays] MASGQES--REELARMAEEGQTVV---ARGGKTLAQEHLAEGRSRSH---GGQTRSEQLGHEGYSEMGSKGGQTRKEQLGHEGYSEMGRK-----
 XhLEA1-2 (T2)X. humilis MASEQE---RAELDQKAREGETVVPGGTGGKSLAQEHLAEGRSR---GGQTRREQLGTEGYKEMSGRGGQTTREQLGTEGYKEMGRK-----
 XhLEA1-1 (M6) X. humilis MASEQE---KAELDKKAREGETVVPGGTGGKSLAQOHLAEGRSR---GGQTRREQLGTEGYKEMGHK-----
 P42755_Em_p_H5_Triticum_aestiv MASGQQE--RSELDRMAREGETVVPGGTGGKSLAQEHLADGRSR---GGQTRKEQLGEEGYREMGRK-----
 CAB59731_Em_H5_protein [Triti MASGQQE--RSELDRMAREGETVVPGGTGGKSLAQEHLADGRSR---GGQTRKEQLGEEGYREMGRK-----
 CAA51700_Em_protein [Triticum MASGQQE--RSQLDRKAREGETVVPGGTGGKSLAEHLAENLAEGRSR---GGQTRREQMGEEGYSEMGRK-----
 CAA68322_Em_protein [Triticum MASGQQE--RSQLDRKAREGETVVPGGTGGKSLAQENLAEGRSR---GGQTRREQMGEEGYSQMGRK-----
 CAA36323_Em [Triticum_aestivum MASGQEK--GRSELDLSLAREGQTVVPGGTGGKSYEAQEKLAEGRSR---GGQTRKEQMGEEGYSEMGRK-----
 NM_129575_AT2G40170_A. thalian MASQQE---KKQLDERAKKGETVVPGGTGGKSFQAQHLAEGRSR---GGQTRKEQLGTEGYQQMGRK-----
 NP_001105429_LEA_EMB564 [Zea m MASGQES--RKELDRKAREGETVVPGGTGGKSVEAQEHLEAEGRSR---GGQTRREQLGQQGYSEMGRK-----
 CAB88086_LEA-polypeptide [Secal MASGQQE--RSQLDRKAREGETVVPGGTGGKSLAQENLAEGRSR---GGQTRREQMGEEGYSEMGRK-----
 P46532_LEA_B19.1B_Hordeum_vulg MASGQQE--RSQLDRKAREGETVVPGGTGGKSLAEHLADNLAEGRSR---GGQTRREQMGEEGYSEMGRK-----
 Q05190_LEA_B19.1A_Hordeum_vulg MASGQQE--RSQLDRKAREGETVVPGGTGGKSLAQOHLAEGRSR---GGQTRREQMGEEGYSEMGRK-----
 NP_001055273_Os05g0349800 [Ory MASGQQQQGRSELDRMAREGQTVVPGGTGGKSLAQENLAEGRSR---GGQTRKEQMGEEGYREMGRK-----
 BAD22768_LEA [Bromus_inermis] MASGQEK--RSQLDRKAREGETVVPGGTGGKSYEAQENLAEGRSR---GGQTRKEQMGEEGYSEMGRK-----
 ABD66069_EMZ08_LEA_oil_palm [E MATRQE---RAELDAKARQGETVVPGGTGGHSLEAQEHLAEGRSR---GGQTRREQLGTEGYQEMGRK-----
 CN202172_Tor2296_Syntrichia_ru -MSGKDA---TQDLDARAAAGETVVPGGTGGKSVEAQOHLAEGRSK---GGQTRAEQLGHDGVYEMGKK-----
 CN202430_Tor2590_Syntrichia_ru -MSGKDA---TQDLDARAAAGETVVPGGTGGKSVEAQOHLAEGRSK---GGQTRAEQLGHDGVYEMGKK-----
 DQ084331_Physcomitrella_patens MSSDDQD-----LDARAAAGETVVPGGTGGKSLAQKNLAEGRSK---GGQTRAEQLGHEGYTEMGKK-----
 Leaf_2_XhLEA1-1S2_X.humilis MASAQE---RIELDRRAREGETVVPGGTGGKSLAQEHLADGRSR---GGQTRRDQLGSEGYSELGRMGGSQSGSYSDEMATGGVLPDLGAGETVVPFG
 Leaf_1_XhLEA1-1S2_consens. X. h MASAQE---RIELDRRAREGETVVPGGTGGKSLAQEHLADGRSR---GGQTRRDQLGSEGYSELGRMGGSQSGSYSDEMATGGVLPDLGAGETVVPFG
 root_XhLEA1-1S2_X. humilis MASAQE---RIELDRRAREGETVVPGGTGGKSLAQEHLADGRSR---GGQTRRDQLGSEGYSELGRMGGSQSGSYSDEMATGGVLPDLGAGETVVPFG
 Seed_XhLEA1-1S2_X. humilis MASAQE---RIELDRRAREGETVVPGGTGGKSLAQEHLADGRSR---GGQTRRDQLGSEGYSELGRMGGSQSGSYSDETATGGVLPDLGAGETVVPFG
 LEA1 motif (PS00431)) -----GETVVPGGTGGKSLAQENLAEGRSR
 LEA1 motif_1 (Tunnacliffe) -----GGQTRREQLGEEGYSQMGRK
 LEA1 motif (PS00431)) -----GETVVP

Degenerate Primers: Forward: TVVPGGT
Reverse: MG (R/H/K) KGG (L/E)

EGYSQMGRK
-----GGQTRRELGEEGYSQMGRK

EGYQEMGHKGGEARKEQLGHEGYKEMGRKGGLSTMEKSGGERAE EEGIEIDESKFTNK-----
EGYSEIGHKGGETRRDQIGSE EYREMGGRKGGLSTKDAESGGERAAREGIEIDESKYKTNV-----
EGYREMGHKGGETRKEGEEGYREMGGRKGGLSTMNESGGERAAREGIDIDESKFKTKS-----
EGYKEMGRKGGETRKEQLGEEGYSEMGGRKGGLSTKDAESGGERAAREGIDIDESKFKTKS-----

EGYQEMGKK-----GGLSTKEESGGERAAREGIEIDESKFRTKS-----
EGYHEMGRK-----GGLSTMEESSGGERAAREGIDIDESKFKTKS-----

GGLSTMQESGGERAAREGIEIDESKFRTKS
 GGLSTMQES-GERAAREGIEIDESKFRTKS
 GGLSTGEESGGERAEREGIEIDESKYKTKS
 GGLSTGEESGGERAEREGIHIDESKYKTKS
 GGLSTMESGGERAAREGIEIDESKFKTKS
 GGLSTMESGGERAAREGIEIDESKFKTKS
 GGLSTNDESGGERAAREGIDIDESKFKTKS
 GGLSTNDESGGDRAAREGIDIDESKFKTKS
 GGLSTNDESGGERAAREGIDIDESKFKTKS
 GGLSTGDKPGGEHAE EEGVEIDESKFRTKT
 GGLSTTDESGGERAAREGVTIDESKFTK
 GGLSTMDESGGERAAREGIDIDESKFKTKS
 GGLSTNDESGGERAAREGIDIDESKFKTKS
 GGLSSNDESGGERAAREGIDIDESKFKTKS
 GGLSTGDESGGERAAREGIDIDESKYKTKS
 GGLSTNDESGGERAAREGIDIDESKFKTKS
 GGLSTTDESGGERAAREGIQIDESKFRT

GGSATNDMSGGEAADAAAGREIDASKFTN
 GGSATNDKSGGEAADAAAGREIDESKFTN
 GGSATNEMSGGEAAFAAGREIDESKFTNQ

GAGGKSVEAQENLARAQLQWMKDAGDAAGRRAGSSSEPKVTVHSAAQGEAAERHLSPTRSPES
GAGGKSVEAQENLARGLLDLLSVTY-----
GAGGKSVEAQENLARGRRRGGETRREQLGTEGYSALGRTGGSS-----
GAGGKSVEAQENLAGRRRGGETRKEQLGAEGYSALGRTGGSS-----

GTGGKSLEAQNLAEGRSR
VPGGT
MG (R/H/K) KGG (L/E)

Appendix E: Immune response of New Zealand white rabbits to GST-*XhLEA1-4S1* and GST-*XhLEA1-1S2* recombinant protein antigens.

Serum Dilution	OD Readings at 405nm using pNPP substrate										
	-Ve Control	GST- <i>XhLEA1-4S1</i> antisera					GST- <i>XhLEA1-1S2</i> antisera				
		Pre-bleed	First bleed	Second bleed	Third bleed	Forth bleed	Pre-bleed	First bleed	Second bleed	Third bleed	Forth bleed
1/10 ¹ *	0.001	0.762	0.935	0.608	0.557	0.617	0.821	0.877	0.819	0.929	0.829
1/10 ²	0.0	0.2	1.598	1.502	1.547	1.47	0.411	1.478	1.465	1.478	1.34
1/10 ³	0.0	0.073	1.486	1.5	1.591	1.626	0.108	1.569	1.569	1.608	1.639
1/10 ⁴	0.0	0.013	0.933	0.916	1.316	1.261	0.022	1.003	0.952	1.094	1.18
1/10 ⁵	0.001	-0.002	0.257	0.316	0.628	0.55	0.003	0.267	0.301	0.547	0.411
1/10 ⁶	0.0	-0.004	0.034	0.168	0.14	0.148	-0.009	0.07	0.185	0.158	0.095
1/10 ⁷	0.0	-0.014	0.043	0.059	0.052	0.029	-0.022	0.079	0.042	0.072	0.041
1/10 ⁸	0.0	0.016	-0.024	0.002	-0.017	-0.02	-0.034	0.027	0.012	0.046	0.028

* Lower OD values for first to fourth bleeds at this dilution was because of higher OD reading of the pre-bleed. The OD of the highest dilution near the cut-off point is indicated in bold.

Appendix F: pET-21a(+) vector map

